



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2021

Verstrickt in der Rhetorik der Technologie-, Daten- und Medienunternehmen? Zu Problematiken im Umgang mit nicht-wissenschaftlichen Nutzungsdaten

Fürst, Silke

Abstract: Die Kommunikations- und Medienwissenschaft sowie benachbarte Disziplinen fragen zunehmend danach, wie soziales Handeln durch Daten bzw. digitale Spuren gemessen, ausgewertet und repräsentiert wird. Es gibt jedoch nur wenige Analysen dazu, wie die Kommunikations- und Medienforschung selbst in unterschiedlichen Formen solche nichtwissenschaftlichen Nutzungsdaten verwendet. Dieser Beitrag zeigt, dass Nutzungsdaten von Technologie-, Daten-, Marktforschungs- und Medienunternehmen vielfach mit starker Intransparenz verbunden sind und von der Forschung oft nicht ausreichend hinterfragt und kontextualisiert werden. Über die Aussagekraft der verschiedenen Metriken und Rankings wird zu wenig kritisch diskutiert und auch die mit ihnen verfolgten kommerziellen Interessen und Strategien werden selten in den Blick genommen. Die vorliegende Analyse identifiziert und diskutiert fünf problematische Umgangsweisen mit nicht-wissenschaftlichen Nutzungsdaten: A) Mangelnde Einordnung der Aussagekraft von Nutzungsmetriken und -rankings, B) Mangelnde Einordnung von Rankings als strategische Unternehmensinstrumente, C) Unbegründete Inferenz von Nutzungsdaten auf Makro-Phänomene, D) Unkritische Übernahme von interessengeleiteten Superlativen und Generalisierungen und E) Verwendung von geschätzten Zahlen als gemessene Daten. Die Analyse dieser Problematiken wirft einen kritischen Blick auf die Kommunikations- und Medienforschung, die zum Teil mangelnde Sicherung von wissenschaftlicher Qualität und jüngste Entwicklungen im Wissenschaftssystem. Sie eröffnet zudem neue Fragen und Perspektiven, mit denen die Kommunikations- und Medienforschung einen wichtigen Beitrag leisten kann. Denn gerade in Gesellschaften, in denen Nutzungszahlen und -rankings omnipräsent sind und mächtige Unternehmen den Datenzugang weitgehend kontrollieren, ist ein kritischer Blick auf Nutzungsdaten unverzichtbar.

DOI: <https://doi.org/10.5771/2192-4007-2021-2-I>

Other titles: Caught up in the rhetoric of technology, data, and media companies? Problematic practices in dealing with nonscientific usage data

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-205033>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:

Fürst, Silke (2021). Verstrickt in der Rhetorik der Technologie-, Daten- und Medienunternehmen? Zu Problematiken im Umgang mit nicht-wissenschaftlichen Nutzungsdaten. *SCM Studies in Communication and Media*, 10(2):175-221.
DOI: <https://doi.org/10.5771/2192-4007-2021-2-I>

FULL PAPER

Verstrickt in der Rhetorik der Technologie-, Daten- und Medienunternehmen?

Zu Problematiken im Umgang mit
nicht-wissenschaftlichen Nutzungsdaten

Caught up in the rhetoric of technology, data, and media companies?

Problematic practices in dealing with nonscientific usage data

Silke Fürst

Silke Fürst (M.A.), IKMZ – Department of Communication and Media Research, University of Zurich, Andreasstrasse 15, 8050 Zurich, Switzerland; Contact: s.fuerst(at)ikmz.uzh.ch. ORCID: <https://orcid.org/0000-0002-1978-4277>

Verstrickt in der Rhetorik der Technologie-, Daten- und Medienunternehmen?

**Zu Problematiken im Umgang mit
nicht-wissenschaftlichen Nutzungsdaten**

Caught up in the rhetoric of technology, data, and media companies?

Problematic practices in dealing with nonscientific usage data

Silke Fürst

Zusammenfassung: Die Kommunikations- und Medienwissenschaft sowie benachbarte Disziplinen fragen zunehmend danach, wie soziales Handeln durch Daten bzw. digitale Spuren gemessen, ausgewertet und repräsentiert wird. Es gibt jedoch nur wenige Analysen dazu, wie die Kommunikations- und Medienforschung selbst in unterschiedlichen Formen solche nicht-wissenschaftlichen Nutzungsdaten verwendet. Dieser Beitrag zeigt, dass Nutzungsdaten von Technologie-, Daten-, Marktforschungs- und Medienunternehmen vielfach mit starker Intransparenz verbunden sind und von der Forschung oft nicht ausreichend hinterfragt und kontextualisiert werden. Über die Aussagekraft der verschiedenen Metriken und Rankings wird zu wenig kritisch diskutiert und auch die mit ihnen verfolgten kommerziellen Interessen und Strategien werden selten in den Blick genommen. Die vorliegende Analyse identifiziert und diskutiert fünf problematische Umgangsweisen mit nicht-wissenschaftlichen Nutzungsdaten: A) Mangelnde Einordnung der Aussagekraft von Nutzungsmetriken und -rankings, B) Mangelnde Einordnung von Rankings als strategische Unternehmensinstrumente, C) Unbegründete Inferenz von Nutzungsdaten auf Makro-Phänomene, D) Unkritische Übernahme von interessengeleiteten Superlativen und Generalisierungen und E) Verwendung von geschätzten Zahlen als gemessene Daten. Die Analyse dieser Problematiken wirft einen kritischen Blick auf die Kommunikations- und Medienforschung, die zum Teil mangelnde Sicherung von wissenschaftlicher Qualität und jüngste Entwicklungen im Wissenschaftssystem. Sie eröffnet zudem neue Fragen und Perspektiven, mit denen die Kommunikations- und Medienforschung einen wichtigen Beitrag leisten kann. Denn gerade in Gesellschaften, in denen Nutzungszahlen und -rankings omnipräsent sind und mächtige Unternehmen den Datenzugang weitgehend kontrollieren, ist ein kritischer Blick auf Nutzungsdaten unverzichtbar.

Schlagwörter: Nutzungszahlen, Datafizierung, Quantifizierung, Unternehmensstrategien, Wissensproduktion.

Abstract: Communication and media researchers, as well as scholars from related disciplines, are increasingly asking how social practices are measured, analyzed, and represent-

ed through data or digital traces. However, there is little research on the various ways in which communication and media research itself uses such nonscientific usage data. This paper shows that usage data from technology, data, market research, and media companies are in many cases opaque and often insufficiently scrutinized and contextualized by scholars. There is too little critical discussion and analysis as to what different metrics and rankings reveal and what commercial interests and strategies are driving them. The analysis at hand identifies and discusses five problematic ways of dealing with nonscientific usage data: A) lacking contextualization of what usage metrics and rankings indicate, B) lacking contextualization of rankings as strategic business tools, C) unsubstantiated inference from usage data to macro phenomena, D) uncritical adoption of superlatives and generalizations created by vested interests, and E) use of estimated numbers as measured data. The analysis of these problematics sheds a critical light on communication and media research, the partial lack of ensuring scientific quality, and recent developments in science. It also opens up new questions and perspectives with which communication and media research can make an important contribution. A critical look at usage data is essential, especially in societies in which usage figures and rankings are omnipresent and powerful corporations largely control access to data.

Keywords: Audience numbers, datafication, quantification, corporate strategies, knowledge production.

1. Einleitung

In der Kommunikations- und Medienwissenschaft und Soziologie haben die Konzepte der Datafizierung und Quantifizierung des Sozialen in den letzten Jahren deutlich an Bedeutung gewonnen (E. P. Berman & Hirschman, 2018; Bolin & Velkova, 2020; Bossetta, 2018; Dencik, 2020; Heintz, 2010; Hepp, 2016; Livingstone, 2019; Mau, 2019; van Dijck & Poell, 2013; Wu & Harsh, 2020). ForscherInnen fragen in diesem Rahmen danach, wie soziales Handeln durch Daten erfasst, repräsentiert und beeinflusst wird. Sie stellen fest, dass Daten zunehmend den Alltag durchdringen „and become part of the social world“ (Beer, 2018, S. 467). Dabei werden insbesondere die mit der Internetnutzung einhergehende automatische Datengenerierung und -verarbeitung in den Blick genommen. Arbeiten zu Datafizierung und Quantifizierung schließen aber auch andere datengetriebene Praktiken ein (Christin, 2020), wie etwa die seit vielen Jahrzehnten etablierte kommerzielle Publikumsforschung (Bolin & Velkova, 2020; Jensen, 2019; Wehner, 2020). Im Kern geht es darum, dass soziales Handeln zunehmend durch Daten repräsentiert wird und die Produktion und Auswertung von Daten als Basis für Wirklichkeitskonstruktionen Akzeptanz findet. Den Anschein von Objektivität erhalten Daten dadurch, dass übergangen oder vergessen wird, wie sie gewonnen wurden und welche Voraussetzungen und Entscheidungen mit ihrer Gewinnung verbunden waren (Beer, 2016b; Heintz, 2010; Mau, 2019; Van Witsen, 2020).

Kommunikations- und MedienforscherInnen sowie SoziologInnen interessieren sich zwar zunehmend dafür, wie Nutzungsdaten die Produktion, Darstellung und Nutzung von Inhalten beeinflussen. Allerdings wendet die Forschung diese Reflexion kaum auf sich selbst an (so auch Waldherr, Geise, & Katzenbach, 2019, S. 3955): Es wird vergleichsweise wenig darüber diskutiert, wie Nutzungsdaten,

die nicht mit wissenschaftlichen Methoden gewonnen wurden, Eingang in die Forschung finden und als Repräsentation sozialer Wirklichkeit akzeptiert werden. Van Dijck (2014, S. 199) hat bereits vor mehreren Jahren problematisiert, dass ForscherInnen Social Media-Daten häufig als „natural traces“ der NutzerInnen behandeln. Inzwischen kritisieren einige AutorInnen die Dominanz von datengetriebenen Social-Media-Analysen und stellen einen Mangel an kritischen Perspektiven fest (Bruns, 2020; Fuchs, 2021, S. 110–117; Jungherr, Schoen, Posegga, & Jürgens, 2017; Lohmeier, 2014; Lomborg, 2017; Marres, 2017; Schmitt, 2018).

Dieser attestierte Mangel bezieht sich auf zwei unterschiedliche Formen von wissenschaftlicher Kritik. Zum einen gilt Kritik, das heißt die Infragestellung, Prüfung und Diskussion von Aussagen und Theorien und deren Geltung bzw. Bewährtheit, als grundlegender Teil wissenschaftlichen Arbeitens und wissenschaftlicher Erkenntnisgewinnung (Merton, 1973, S. 264–265; Popper, 2017 [1962], S. 217, 221; 1984, S. 18, 224–225; Prim & Tilmann, 1997, S. 10). Die kritische Betrachtung empirischer Aussagen und theoretischer Annahmen wird damit als Teil der Wissenschaftslogik verstanden. Zum anderen kann sich Kritik auch auf die Rolle der Wissenschaft beziehen, das heißt auf das Verhältnis zwischen Wissenschaft und Gesellschaft. Kritische Wissenschaft in diesem Verständnis meint, dass die Auswahl von Forschungsfragen auch davon geleitet wird, gesellschaftliche Machtverhältnisse und Ungleichheiten aufzudecken, zu problematisieren und damit Gesellschaftskritik zu üben. Sie zielt darauf, nicht den Status quo zu zementieren, sondern durch das Aufdecken gesellschaftlicher Probleme einen Beitrag dazu zu leisten, dass diese adressiert, öffentlich diskutiert und abgebaut werden können (Drori, 2020; Grittmann & Drüeke, 2016; Iliadis & Russo, 2016). Mit Blick auf die Kommunikations- und Medienforschung bedeutet dies etwa, strategische Einflussnahmen auf die öffentliche Kommunikation zu untersuchen und Zusammenhänge von Medienstrukturen, Nutzungsdaten, Macht und Ungleichheit zu problematisieren (Dencik, 2020; Krüger & Sevnigani, 2019; Lewis & Molyneux, 2018, S. 18–19; Splichal, 2008).

Diese beiden Formen von Kritik sind bisher in Bezug auf den Umgang mit nicht-wissenschaftlichen Nutzungsdaten vergleichsweise wenig geübt worden. Bisher mangelt es an umfassenden Analysen dazu, wie die Kommunikations- und Medienwissenschaft sowie benachbarte Disziplinen mit jenen Daten umgehen, die durch Technologie-, Daten-, Beratungs- und Medienunternehmen generiert und veröffentlicht werden (siehe aber kürzlich: Wu & Harsh, 2020). Dies ist nicht allein im Kontext von Digitalisierung und Big Data wichtig, sondern betrifft auch Nutzungszahlen zu traditionellen Medien wie Rundfunk und Presse.

Dieser Beitrag hat zum Ziel, problematische Praktiken zu identifizieren und zu beleuchten und damit zu einer Sensibilisierung im Umgang mit nicht-wissenschaftlichen Nutzungsdaten beizutragen. Im Fokus stehen damit wissenschaftliche Standards der Wissensproduktion und die Frage, inwieweit diese Standards durch eine zu leichtfertige Übernahme von nicht-wissenschaftlichen Daten und unternehmerischer Rhetorik gefährdet werden. Dazu wird zunächst eingeordnet, wie die Frage nach dem Umgang mit nicht-wissenschaftlichen Nutzungsdaten angesichts von Wissenschafts- und Medienwandel an Dringlichkeit gewinnt. In diesem Rahmen wird auch dargelegt, welche Bedeutung Zahlen und Daten in Wissen-

schaft und Gesellschaft haben (Abschnitt 2). Im Anschluss werden fünf Problematiken im Umgang mit nicht-wissenschaftlichen Nutzungsdaten herausgearbeitet und an Beispielen diskutiert (Abschnitt 3). Im Fazit wird schließlich zusammenfassend diskutiert, was dies für die Kommunikations- und Medienwissenschaft bedeutet und welche positiven Potenziale aus einer solchen Kritik gewonnen werden können.

2. Die Bedeutung von Zahlen im Kontext von Wissenschafts- und Medienwandel

Dieser Beitrag geht der Frage nach, in welchen Formen nicht-wissenschaftliche Nutzungsdaten Eingang in die wissenschaftliche Forschung zu Medien und Kommunikation finden. Der Fokus liegt dabei auf problematischen Praktiken. Erst durch die Beleuchtung problematischer Praktiken kann ein Problembewusstsein dafür gewonnen werden, inwiefern wichtige Standards der wissenschaftlichen Erkenntnisproduktion nicht ausreichend funktionieren. Mit nicht-wissenschaftlichen Nutzungsdaten sind jene Zahlen gemeint, die Technologie-, Daten-, Marktforschungs-, Beratungs- und Medienunternehmen bezüglich der Mediennutzung und -verbreitung kommunizieren. Bedeutsam ist, dass diese Zahlen in einem kommerziellen Kontext entstehen und ihre Kommunikation einem unternehmerischen Eigeninteresse folgt.

Für die hier behandelte Problematik sind Strukturen und Wandlungsprozesse im Wissenschafts- und Mediensystem und die Bedeutung von Daten in Wissenschaft und Gesellschaft relevant. Betrachten wir zunächst die Wissenschaft, so ist ein Ineinandergreifen mehrerer Entwicklungen zu beobachten. Zum einen hat sich der Publikationsdruck und -rhythmus in den vergangenen Jahren stark erhöht (Glerup & Horst, 2014; Probst, Buhmann, Ingenhoff, & Lepori, 2019; Taubert & Weingart, 2016). Manche Kommunikations- und Medienforscher sprechen gar von einem „hyperactive“ pace of publishing“, der die Generierung von „quick hit studies“ befördere (Lewis & Molyneux, 2018, S. 18). Es gibt verstärkt Anreize, jene Themen zu behandeln, die einen hohen und stetigen „Output“ (Ignatow, 2020, S. 5–6) versprechen. Damit werden Forschungsfragen attraktiv, die sich möglichst schnell bearbeiten lassen, die den aktuellen Forschungstrends folgen und zu denen Daten relativ leicht verfügbar sind. Zum anderen hat sich im Zuge der Digitalisierung der Zugang zu großen Datensätzen enorm vereinfacht. Immer häufiger steht nun am Beginn der Forschung nicht eine Forschungsfrage, die im Rahmen eines Entdeckungszusammenhangs besonders relevant erscheint (wie es als wissenschaftlicher Standard formuliert wird: Brosius, Haas, & Koschel, 2016, S. 28; Scheufele & Engelmann, 2009, S. 31; Tausendpfund, 2018, S. 35–37), sondern die Verfügbarkeit von Daten und großer Datensätze. Viele WissenschaftlerInnen beobachten insbesondere im Bereich der Forschung zu Social Media im Allgemeinen und Twitter im Speziellen eine Dominanz von pragmatischen, datengetriebenen Vorgehensweisen, die Kontexte und konzeptionelle Fragen vernachlässigen (Bruns, 2019, S. 1545; Freelon, 2014, S. 69–70; Fuchs, 2021, S. 110–113; Jungherr et al., 2017, S. 351–352; Lohmeier, 2014, S. 77, 85–86;

Lomborg, 2017, S. 8; van Dijck, 2014, S. 199; Waldherr et al., 2019, S. 3955, 3967). Marres sieht darin die „troubling tendency to go with whatever ontology or methodology happens to be hard-wired in the apparatus that generates the data“ (Marres, 2017, S. 187; siehe auch Driscoll & Walker, 2014).

Auch das Mediensystem hat sich in den vergangenen Jahrzehnten stark gewandelt und ausdifferenziert. Immer mehr Medienkanäle, Plattformen und Inhalte konkurrieren um die begrenzte Aufmerksamkeit der NutzerInnen (Jarren, 2015; Jarren, Künzler, & Puppis, 2019; Stark, 2013; Webster, 2014). Durch diese zunehmende Konkurrenz erhält die Bindung von starker Aufmerksamkeit einen besonderen Wert und wird innerhalb des Mediensystems in vielfachen Formen als „Popularität“ kommuniziert (Fürst, 2014, 2021; Fürst & Oehmer, 2021; Gillespie, 2016; Haim, Kümpel, & Brosius, 2018; Webster, 2014). Wer mittels Nutzungszahlen kommunizieren kann, dass er viel Aufmerksamkeit erregt oder gar im Mittelpunkt kollektiver Aufmerksamkeit steht, kann mit weiterer Aufmerksamkeit und journalistischer Berichterstattung rechnen (Fürst, 2020; Fürst & Oehmer, 2021; Stauff & Thiele, 2007). Die großen Technologie- und Datenunternehmen waren in den vergangenen Jahren sehr bestrebt, ihre jeweiligen Nutzungsdaten und -metriken öffentlich sichtbar zu machen und durch Medienaufmerksamkeit bedeutsam erscheinen zu lassen (Anstead & O’Loughlin, 2015; Baldwin-Philippi, 2019; Kreiss & McGregor, 2018; McGregor, 2019). Mit den Worten von van Dijck und Poell (2013, S. 7): „Each corporation actively seeks to promote their popularity and ranking mechanisms in order to enhance the value of its platform and its users.“ Da hier unternehmerische Eigeninteressen im Mittelpunkt stehen, findet diese Kommunikation von Nutzungszahlen selektiv statt und ist oft mit starker Intransparenz verbunden (Anstead & O’Loughlin, 2015; Fürst, 2021; Gillespie, 2011; Wu & Harsh, 2020). Damit wird eine stetige Medienberichterstattung über Nutzungsdaten stimuliert. Der Journalismus hat generell eine starke Affinität zu Zahlen (Harrison, 2021; Lawson, 2021; Luhmann, 1997, S. 1099–1100). Es ist die komprimierte Form und scheinbare Objektivität von Zahlen, die diese für den Journalismus attraktiv machen (Curtin & Maier, 2001; Heintz, 2010; Roeh & Feldman, 1984). Mehrere Studien haben gezeigt, dass JournalistInnen häufig das Zahlenmaterial aus Medienmitteilungen übernehmen, ohne dieses zu hinterfragen und zu kontextualisieren (Ahmad, 2016; Harrison, 2021; Lawson, 2021; Nguyen, 2018; Van Witsen, 2020). Auch JournalistInnen selbst nehmen wahr, dass sie Zahlen weitgehend unhinterfragt übernehmen und PR-Quellen mittels Zahlen eine starke Definitionsmacht gewinnen können (Curtin & Maier, 2001). Diese Problematik nimmt durch die enormen redaktionellen Stellen- und Ressourcenkürzungen im vergangenen Jahrzehnt potenziell noch zu (Fürst, Meißner, Hofstetter, Puppis, & Schönhagen, 2017; Lawson, 2021; Pickard, 2020).

Die kulturelle Bedeutung von Zahlen hat in den vergangenen Jahren vermehrt die Aufmerksamkeit der Forschung auf sich gezogen. Soziologische Arbeiten haben herausgestellt, dass das Vertrauen in Zahlen „as transparent and objective measurements of the ‚real‘ world“ (Christin, 2020, S. 117) eine lange Geschichte hat. Seit der im 18. Jahrhundert einsetzenden amtlichen Statistik haben sich Daten zunehmend als Form der gesellschaftlichen Selbstbeschreibung etabliert. Obschon Daten stets im Kontext ihrer Generierung betrachtet und interpretiert wer-

den müssen (boyd & Crawford, 2012), sind sie kulturell mit einer spezifischen Legitimation verbunden: Zahlen *gelten* gemeinhin als neutral und präzise und werden oftmals mit Objektivität und Wissenschaftlichkeit assoziiert (Beer, 2016b; Diaz-Bone, 2018; Heintz, 2010; Mau, 2019; Porter, 1996; Prietl & Houben, 2018; Splichal, 2008; Van Witsen, 2020). Auch Big Data sind diskursiv mit dem Mythos verbunden worden, dass sie einen objektiven und neutralen Zugang zur komplexen Wirklichkeit bieten würden (Beer, 2016a; boyd & Crawford, 2012, S. 663). Unternehmen, die digitale Datenanalysen anbieten, knüpfen an diese Bedeutungszuweisungen an und betonen, dass sie die Wirklichkeit mittels Daten allumfassend und präzise abbilden können (Beer, 2018). Durch die Digitalisierung und die Marktmacht der großen Technologiekonzerne (van Dijck, Poell, & de Waal, 2018) sind Nutzungsdaten und -rankings in der öffentlichen Kommunikation inzwischen omnipräsent und beeinflussen, was besonders sichtbar wird und bedeutsam erscheint. Die mittels Zahlen typischerweise angestellten Vergleiche (Heintz, 2010) kommen im Wettbewerb um Aufmerksamkeit zum Tragen (Webster, 2014), implizieren Sieger und Verlierer (Beer, 2016b, S. 17–26; Fürst, 2021; Heintz, 2019) und signalisieren, etwas über die Art der Öffentlichkeit und die soziale Bedeutsamkeit eines Inhaltes aussagen zu können (Bolin & Velkova, 2020; Fürst & Oehmer, 2021; Gillespie, 2016, S. 67; Jarren, 2015; Mau, 2019; Wehner, 2020; Winter, Metzger, & Flanagin, 2016). Zudem sind Nutzungszahlen, die von der kommerziellen Publikumsforschung und von Medien-, Technologie- und Datenunternehmen kommuniziert werden, bereits seit vielen Jahren ein so geläufiger Bestandteil des öffentlichen Diskurses, dass auch WissenschaftlerInnen in der Regel unterstellen, dass diese Zahlen durch Messungen entstanden sind und das abbilden, was sie vorgeben. Tatsächlich kursieren aber, wie im Weiteren genauer gezeigt wird, viele Nutzungszahlen, deren Entstehung intransparent und Bedeutung unklar ist oder die wider dem ersten Anschein nicht auf einem Messvorgang beruhen.

Eine kritische Auseinandersetzung zum wissenschaftlichen Umgang mit kommerziellen Nutzungsdaten ist daher wichtig. Die Zunahme von datengetriebenen „quick hit studies“ (Lewis & Molyneux, 2018, S. 18) kann etablierte wissenschaftliche Qualitätsstandards langfristig aushöhlen. Die Übernahme und Einbindung von Zahlen, die sich nicht überprüfen lassen und/oder deren Generierung intransparent ist, steht im Gegensatz zum methodengeleiteten Vorgehen als Grundprinzip der Sozialwissenschaften. Empirische Sozialforschung folgt anerkannten Regeln der Generierung und Auswertung empirischer Daten (z. B. Häder, 2019). Die „Befolgung eines Regelkanons“ (Brosius et al., 2016, S. 9) dient der Qualitätssicherung der Daten sowie der intersubjektiven Nachvollziehbarkeit der Erkenntnisse. Entsprechend spielen die Entwicklung und Vermittlung dieser Standards in der Kommunikations- und Medienwissenschaft (Matthes et al., 2011; Schönhagen & Trebbe, 2005, S. 42) sowie in den Sozialwissenschaften im Allgemeinen eine große Rolle. In Grundlagenbüchern zu Wissenschaftstheorie und sozialwissenschaftlicher Methodik wird in der Regel davon ausgegangen, dass sich Wissenschaft auf Daten stützt, die sie selbst systematisch generiert hat (Brosius et al., 2016; Häder, 2019; Opp, 2014; Prim & Tilmann, 1997, S. 11; Scheufele & Engelmann, 2009; Steininger & Hummel, 2015) und im Rahmen von Primär-

oder Sekundäranalysen auswertet (Baur & Blasius, 2019; Tausendpfund, 2018). Zudem wurde die auf quantitative Forschung ausgerichtete, „empirisch-sozialwissenschaftliche Wende“ der Kommunikationswissenschaft (Löblich, 2010) auch dadurch vorangetrieben, dass Medienpolitik und Medienorganisationen Bedarf nach solchen wissenschaftlichen Daten signalisierten. Umso problematischer wäre es, wenn sich die Kommunikations- und Medienforschung nun abseits von etablierten sozialwissenschaftlichen Methoden zunehmend an verfügbaren Nutzungsdaten orientieren würde, ohne diese ausreichend zu hinterfragen und zu kontextualisieren. Dies widerspräche auch ethischen Grundsätzen. Mit dem 2015 verabschiedeten Ethik-Kodex der DGPK (Filipović, Klaus, & Strippel, 2016) verpflichten sich die Mitglieder „zum Grundsatz der inhaltlichen und methodischen Transparenz ihrer Arbeit“ und dazu, „Einzelheiten der zugrunde liegenden Theorien, Methoden und des Forschungsdesigns [offenzulegen], die für die Beurteilung der Forschungsergebnisse und der Grenzen ihrer Gültigkeit wichtig sind“ (siehe <https://tinyurl.com/dgpuk-ethik>). Deshalb ist es wichtig, problematische Praktiken im Umgang mit nicht-wissenschaftlichen Nutzungsdaten zu identifizieren und zu kritisieren.

3. Problematiken im Umgang mit nicht-wissenschaftlichen Nutzungsdaten

In der kommunikations- und medienwissenschaftlichen sowie soziologischen Literatur gibt es kritische Reflexionen zum Umgang mit nicht-wissenschaftlichen Nutzungsdaten hinsichtlich spezifischer Problematiken oder einzelner Forschungsfelder. Diese Reflexionen sind wichtige Schritte hin zu einem stärkeren Problembewusstsein. Angesichts der Bedeutung, die nicht-wissenschaftliche Nutzungsdaten im Zuge der Digitalisierung und des „computational turn“ (Waldherr et al., 2019, S. 3955) in der Kommunikations- und Medienforschung gewonnen haben, sind mehr Anstrengungen in diesem Bereich notwendig. Derzeit mangelt es an Analysen, die verschiedene Problematiken in unterschiedlichen Forschungsbereichen identifizieren, zusammenführen und diskutieren. Indem wir nicht allein auf einzelne Forschungsbereiche fokussieren, sondern übergreifende Problematiken und Muster herausstellen, wird es möglich, eine Reflexion zum Umgang mit nicht-wissenschaftlichen Nutzungsdaten im gesamten Fach anzustossen und darüber zu diskutieren, wie die Einhaltung wissenschaftlicher Standards zukünftig stärker gesichert werden kann. Dieser Beitrag unternimmt eine solche Analyse und stellt fünf grundlegende Problematiken (A – E) im Umgang mit nicht-wissenschaftlichen Nutzungsdaten heraus und diskutiert diese an Hand von Beispielen aus zahlreichen Forschungsbereichen. Dabei werden sowohl Daten zur Nutzung von Onlinemedien als auch von Rundfunk und Presse berücksichtigt.

Mit Problematiken oder problematischen Praktiken ist hier gemeint, dass die zentralen Gütekriterien der Reliabilität und/oder Validität (Brosius et al., 2016, S. 51–58; Häder, 2019, S. 109–116) von Daten und Aussagen nicht ausreichend gegeben sind bzw. nicht überprüft werden können. Problematische Praktiken werden in Büchern zu Methoden und Methodologie der Sozialwissenschaften behandelt, damit diese leichter vermieden werden können (siehe z. B. Brosius et al., 2016, S. 51–55, 68, 121; Häder, 2019, S. 121–122, 326; Opp, 2014, S. 73–74). In

der hier vorliegenden Analyse werden auch problematische Praktiken berücksichtigt, in denen Nutzungszahlen verbreitet werden, die nicht auf einem Messvorgang beruhen. Es handelt sich hierbei also um empirische Sätze mit deskriptiven Aussagen, die ihrer Form nach eine aposteriorische Gültigkeit haben, also mit „Erfahrungen der Realität konfrontiert worden“ (Prim & Tilmann, 1997, S. 62) sein müssen – tatsächlich aber a priori getroffen wurden (vgl. Prim & Tilmann, 1997, S. 60–73).

A. Mangelnde Einordnung der Aussagekraft von Nutzungsmetriken und -rankings

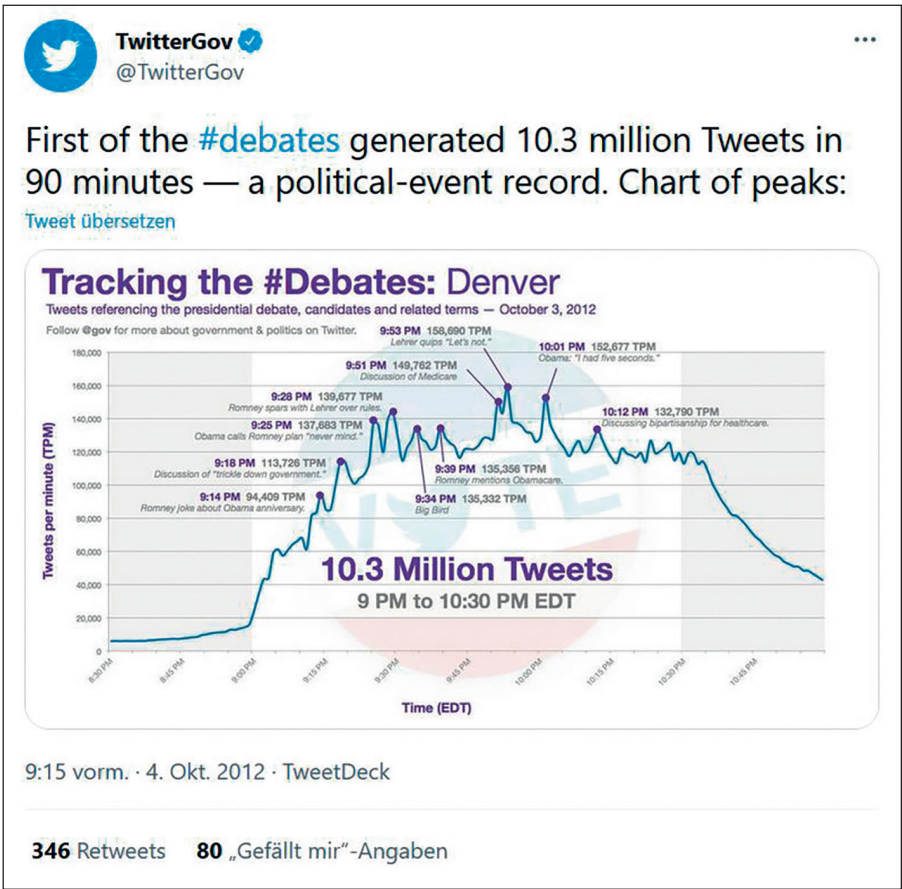
Mit der Etablierung neuer Plattformen und Online-Nachrichtenseiten haben sich auch neue Nutzungsmetriken (z. B. Follower, Likes, Shares) und Nutzungsrankings (z. B. meistgelesene Artikel) etabliert. Das Geschäftsmodell der Plattformen basiert auf der Generierung von massiven Nutzungsdaten (Bakardjieva & Gehl, 2017; Bossetta, 2018; van Dijck, 2013, 2014). Gleichzeitig werden Metriken und Rankings als „popularity cues“ (Haim et al., 2018; Winter et al., 2016, S. 683) für die NutzerInnen selbst sichtbar gemacht. Damit sollen wiederum weitere Nutzungsaktivitäten stimuliert werden, die sich monetarisieren lassen (van Dijck, 2013; Waldherr et al., 2019, S. 3960; Webster, 2014; Wu & Harsh, 2020). Angesichts dieser ökonomisch-strategischen Bedeutung ist es problematisch, dass die bisherige Forschung vergleichsweise wenig die Infrastrukturen der Plattformen untersucht hat sowie die Narrative und rhetorischen Mittel, mit denen Technologie-, Medien- und Datenunternehmen ihre Nutzungsmetriken und -daten kulturell etablieren (Beer, 2018, S. 466; Bossetta, 2018, S. 472; Gerlitz & Rieder, 2018, S. 529; Lewis & Molyneux, 2018, S. 18; Turow & Couldry, 2018, S. 419). In einer Art „data rush“ (Mahrt & Scharkow, 2013, S. 22) bzw. „‘big data’ gold-rush“ (Bruns, 2020, S. 77) lag der Fokus stattdessen insbesondere auf jenen Metriken, die die Plattformen zur Datenstimulierung einsetzen und selbst definieren. Die Omnipräsenz dieser Metriken in der öffentlichen Kommunikation (Anstead & O’Loughlin, 2015; Fürst & Oehmer, 2021; McGregor, 2019) trug offenbar mit dazu bei, dass in wissenschaftlichen Social Media Analysen häufig kaum geklärt oder hinterfragt wurde, was jene Einheiten, die als Follower, Likes, Shares u. ä. gelabelt sind, jeweils genau bedeuten und was deren Anzahl indizieren kann (Freelon, 2014). Zudem übernahmen einige ForscherInnen die von den Plattformen etablierte Unterstellung, dass die dahinterstehenden Medienpraktiken eine bestimmte, praktisch homogene Bedeutung hätten – was der Komplexität und Unterschiedlichkeit des Nutzungsverhaltens nicht entspricht (dazu kritisch Driscoll & Walker, 2014, S. 1747; Freelon, 2014, S. 70; Gerlitz & Rieder, 2018, S. 529–530). Beispielsweise vollziehen NutzerInnen mit dem Like-Button unterschiedliche Handlungen und Strategien und verbinden damit unterschiedliche Bedeutungen (Bucher, 2017; F. Martin & Dwyer, 2019, S. 74–75). Auch die schiefe Anzahl der Follower kann nicht valide mit dem Einfluss oder der Anerkennung eines Akteurs gleichgesetzt werden (Ausserhofer & Maireder, 2013; Freelon, 2014).

Um das Potenzial von Online-Nutzungsdaten für die sozialwissenschaftliche Forschung zu nutzen, bräuchte es verstärkt konzeptionelle Arbeiten und Studien, die Metriken wie Follower, Likes und Shares mit weiteren, auch qualitativen Daten in Beziehung setzen (Freelon, 2014, S. 70; Hepp, Breiter, & Friemel, 2018, S. 443; Wu & Harsh, 2020). Darüber hinaus ist es problematisch, Nutzungsdaten und -rankings pauschal als digitale Spuren von NutzerInnen zu interpretieren und daran die Popularität von Akteuren festzumachen (als Beispiel: Francia, 2018; Pedro-Carañana, Alvarado-Vivas, & López-López, 2020; Van Aelst, van Erkel, D'heer, & Harder, 2017). Denn je stärker bestimmte Metriken sich in der Öffentlichkeit als Ausweis von Erfolg und Status sowie als Werbewährung etablieren, umso mehr nehmen auch die Versuche zu, diese Metriken mittels Zahlungen, gezielter Synchronisierung zahlreicher Accounts und Automatisierungsprozessen zu beeinflussen und zu manipulieren (Boone, Secci, & Gallant, 2018; Fürst, 2021; Gorwa & Guilbeault, 2020; Hopke, 2015; Karpf, 2012, S. 650; Marres, 2017, S. 72–73; Muhle, 2020; Santini, Salles, Tucci, Ferreira, & Graef, 2020; Turow & Couldry, 2018, S. 418; Woolley & Howard, 2019). Das kann auch ein wichtiger Teil des Geschäftsmodells von Plattformen sein (Bossetta, 2018): Beispielsweise warb Twitter bei politischen Akteuren explizit damit, dass diese auf Nutzungsmetriken und -rankings Einfluss nehmen können und Twitter dafür gegen Bezahlung Instrumente bereithält (Twitter, 2014; siehe z. B. auch die aktuellen Informationen zu sogenannten „trend takeovers“, Twitter, 2021). Um die Menge von Accounts sowie Nutzungsaktivitäten und -daten zu stimulieren, hat Twitter auch „eine ideale Umgebung für die Automatisierung von Accounts“ (Muhle, 2020, S. 48–49) geschaffen – wenngleich unter politischem Druck Anpassungen der „rules on the appropriate use of bots and automation“ (Twitter, 2018) versprochen werden und vereinzelt Lösungen von Social Bots stattfinden (Gorwa & Guilbeault, 2020, S. 238, 241; Keller & Klinger, 2021, S. 43–44).

B. Mangelnde Einordnung von Rankings als strategische Unternehmensinstrumente

Weil die Generierung, Aggregation und weitere Stimulierung von Nutzungsdaten von zentraler Bedeutung für Plattformen ist, haben diese ein starkes Interesse daran, ihren Nutzungsmetriken auch außerhalb der eigenen Plattform Bedeutung und Akzeptanz zu verleihen (van Dijck & Poell, 2013, S. 7). Beispielsweise stellt sich Twitter öffentlich als „echo chamber of people's opinions“ dar (van Dijck & Poell, 2013, S. 9) oder als „real-time measure of public opinion“ (Twitter, 2014, S. 11; siehe auch Kreiss & McGregor, 2018, S. 165). Zugleich entwickelt das Unternehmen verschiedene Nutzungsrankings, um damit Aufmerksamkeit in den Medien zu gewinnen (Fürst, 2021; Kreiss & McGregor, 2018, S. 165). Solche Narrative und Strategien sind bisher selten von der Kommunikations- und Medienforschung untersucht worden. Stattdessen hat ein großer Teil der Forschung eher die Legitimation und Bedeutsamkeit von Plattformen unterstützt (Bruns, 2020, S. 75–76; siehe oben die Ausführungen unter A) und – ohne genauere Prüfung – den auf Öffentlichkeitswirksamkeit zielenden Rankings den Status wissenschaftlicher Fakten verliehen.

Abbildung 1. Visualisierung der “conversation” auf Twitter und der “top tweeted moments” (Screenshot eines Tweets von Twitter)



Anmerkungen: Diese auf Twitter gepostete Graphik (<https://twitter.com/TwitterGov/status/253755175460827138>) wurde auch im Twitter Blog veröffentlicht (siehe Sharp, 2012) und stimulierte journalistische Berichterstattung (mit Verweis auf einen Bericht aus *USA Today*: McKinney, Houston, & Hawthorne, 2014, S. 557). Wissenschaftliche Arbeiten übernahmen Teile dieser Daten (z. B. McKinney et al., 2014, S. 557; Peifer & Holbert, 2013, S. 292) oder verwendeten die ganze Graphik (Schill & Kirk, 2015, S. 212), um Aussagen über die Aktivität und die Diskussionen von NutzerInnen auf Twitter zu treffen.

Dazu gehören jene Rankings, die das Unternehmen Twitter auf seinem Blog sowie auf Twitter selbst im Kontext großer, vorab geplanter Ereignisse veröffentlicht hat (z. B. Fernsehdebatten im US-Wahlkampf, Fußballweltmeisterschaften), wie etwa die „top tweeted moments“ und der „share of conversation“ (Fürst, 2021, S. 349–352). Diese Rankings wurden – während oder unmittelbar nach den Ereignissen – auch zahlreichen Nachrichtenmedien zugespielt und sollten so die Berichterstattung über die jeweiligen Ereignisse beeinflussen (Kreiss & McGregor, 2018; McGregor, 2019). Twitter legte allerdings nicht offen, wie die zugrunde liegenden Daten ausgewählt und ausgewertet wurden. Auch veröffentlichte Twit-

ter keine Beiträge, die die Bedeutung dieser genannten Rankings genauer erklären würden. Mit den Daten wurde aber der Anspruch verknüpft, dass jeweils Aussagen, Themen oder Personen abgebildet werden, die während eines Ereignisses auf Twitter am meisten Austausch („discussion“, „conversation“) ausgelöst haben (Sharp, 2012).¹ Dabei wurden auch Zusammenhänge signalisiert, die nicht gemessen wurden. Im Fall der „top tweeted moments“ wurden zwei separate Datenreihen, namentlich die Anzahl der Tweets pro Minute (TpM) und die je Minute auffällig erscheinenden Aussagen und Ereignisse der Fernsehübertragung, miteinander in eine direkte Beziehung gesetzt (siehe Abb. 1 als Beispiel). Die drei stärksten Peaks (hier u. a.: Moderator „Lehrer quips ‚Let’s not‘“) kommunizierte Twitter als meistdiskutierte Momente (Sharp, 2012) bzw. als „most“ oder „top tweeted moments“.²

Die schiere Anzahl der TpM, die sich auf das Fernsehereignis als Ganzes beziehen, erscheint so als eine unmittelbare Reaktion auf einen bestimmten Moment. Dieser Korrelierung liegt die – von Twitter nicht thematisierte – Unterstellung zugrunde, dass NutzerInnen beim „second screening“ (Gil de Zúñiga, Garcia-Perdomo, & McGregor, 2015) stets innerhalb weniger Sekunden reagieren und sich in ihren Tweets und Retweets nur auf das beziehen, was sich unmittelbar zuvor auf dem Bildschirm abgespielt hat.³

Obschon solche Daten auf Grund ihrer Intransparenz und mangelnden Validität offenkundig gegen wissenschaftliche Standards verstoßen, werden sie in wissenschaftlichen Arbeiten als Nutzungsdaten verwendet. Twitter-Daten zum „share of conversation“ (Chadwick & Stromer-Galley, 2016, S. 284), den „most tweeted topics“ (Schill & Kirk, 2015, S. 215) und „top tweeted moments“ (Burk & Grimmer, 2017, S. 95–96; Camaj & Northup, 2019, S. 34; Gainous & Wagner, 2014, S. 53; Gottfried, Hardy, Holbert, Winneg, & Jamieson, 2017, S. 172; Houston, McKinney, Hawthorne, & Spialek, 2013; Schill & Kirk, 2015, S. 211–214; Schroeder, 2016, S. 280–281) werden zitiert, ausführlich wiedergegeben oder sogar graphisch dargestellt, um ein bestimmtes Argument empirisch zu belegen oder ein Ereignis aus der Nutzerperspektive zu beleuchten. Dabei werden die Narrative übernommen, die Twitter entwickelt hat, um außerhalb der eigenen Plattform bzw. in der breiten Öffentlichkeit als Barometer der öffentlichen Meinung wahrgenommen zu werden. Zum Beispiel: „Twitter has allowed a much broader view-ertariat to judge the debate’s defining moments“ (Freelon & Karpf, 2015, S. 394).

- 1 Twitter (2010, 2014, 2021) sowie ein großer Teil der Forschung zu Twitter verwenden die Begriffe „discussion“ und „conversation“ auf Twitter in einem sehr weiten und unscharfen Sinne (z. B. Chadwick & Stromer-Galley, 2016, S. 284; Marchetti & Ceccobelli, 2016, S. 626–640; Schroeder, 2016, S. 279–283; Zheng & Shahin, 2020, S. 338, 341, 343). Mit diesen Begriffen sind nicht dialogische Sequenzen und wechselseitige Bezüge von Akteuren gemeint, wie sonst üblich (Fürst & Schönhagen, 2020), sondern jedwede Art von Aktivität auf Twitter, wie z. B. Tweets, Likes und Retweets ohne Kommentare (dazu kritisch Muhle, Ackland, & Graham, 2018).
- 2 Für die Präsidentschaftsdebatten in 2012 siehe etwa <https://twitter.com/TwitterGov/status/260578102575497216> und <https://twitter.com/TwitterGov/status/260585717418315777>; für 2016 siehe beispielsweise <https://twitter.com/TwitterGov/status/780599941093855232> und <https://twitter.com/TwitterGov/status/785309872829104129>.
- 3 Entsprechend dürfte es beispielsweise keine Tweets geben, die länger als eine Minute retweetet werden und auch keine Inhalte, die sich in genereller Weise über die dargestellten Personen oder das Ereignis äußern.

Manche Studien (Freelon & Karpf, 2015) untersuchen mit eigenen Datenanalysen die von Twitter ausgewiesenen „top tweeted moments“ und verorten ihre Forschungsfrage und ihr gesamtes Untersuchungsdesign im Rahmen dieses intransparenten Rankings. Alle genannten Arbeiten rekurrieren dabei auf Blogbeiträge und Tweets von Twitter oder auf Medienberichte, die diese Twitter-Daten – wie von Twitter intendiert – zum Thema machten. In diesen Studien wird an keiner Stelle angesprochen, dass die zitierten Daten intransparent sind und im Rahmen spezifischer Unternehmensstrategien öffentlich kommuniziert wurden.

Diese Problematik trifft im Kern auch auf viele Studien zu den „trending topics“ zu, die Teil der Nutzeroberfläche von Twitter sind. Einerseits kommuniziert Twitter (2020), dass diese Trends die „beliebtesten Themen auf Twitter zu einem bestimmten Zeitpunkt“ darstellen. Andererseits gesteht das Unternehmen ein, dass es sich bei diesem Ranking nicht um jene Themen handelt, über die am meisten getweetet wird (Twitter, 2010). Die starke, aber zugleich konstante und längerfristige Verwendung eines Hashtags verspricht weniger Chancen auf eine Platzierung in den „trending topics“ als die starke und kurzfristige Verbreitung eines neu gebildeten Hashtags (Gillespie, 2011; Proferes & Summers, 2019). Genauere Informationen zum Algorithmus macht Twitter nicht öffentlich (Gillespie, 2011; Wu & Harsh, 2020). Auch ist intransparent, auf wie vielen Tweets die jeweiligen „trending topics“ basieren, in welchem Verhältnis dies zur gesamten Aktivität auf Twitter steht (Bruns, 2020, S. 75–76) und wie viele NutzerInnen täglich aktiv sind (Bossetta, 2018, S. 481).

Für eine kritisch inspirierte Forschung wäre angesichts dieser Umstände vor allem interessant, Erkenntnisse über den Algorithmus selbst und seine Einflussdynamiken zu gewinnen (siehe Kitchin, 2017 sowie die einleitenden Ausführungen zur Rolle der Wissenschaft bzw. Gesellschaftskritik). Das schließt auch Fragen danach ein, wie die ausgewiesenen Trends von NutzerInnen verstanden, von nach Einfluss strebenden Akteuren und Organisationen antizipiert sowie von JournalistInnen in der Berichterstattung dargestellt werden (Bucher, 2017; Fürst, 2021; McGregor, 2019; Proferes, 2017; Proferes & Summers, 2019). Stattdessen überwiegen eher Studien, die das von Twitter gesetzte Narrativ der „beliebtesten Themen“ übernehmen und die Trends als die „most popular issues in the community“ oder „list of the most tweeted topics“ verstehen (Marchetti & Ceccobelli, 2016, S. 629; sehr ähnlich auch Bastos, Raimundo, & Travitzki, 2013, S. 264; Deller, 2011, S. 219; Florini, 2014, S. 225; Kwak, Lee, Park, & Moon, 2010, S. 592). Damit begründen diese Studien, warum sie dekriptive Datenanalysen zu den „trending topics“ durchführen oder bestimmte Tweets in das Untersuchungsmaterial aufnehmen.

C. Unbegründete Inferenz von Nutzungsdaten auf Makro-Phänomene

In den vergangenen Jahren haben zahlreiche Studien die von den großen Technologieunternehmen etablierten Nutzungsmetriken zum Ausgangspunkt für Big Data Analysen genommen und damit Schlüsse auf Makrophänomene gezogen. So wurden etwa Daten zu Likes, Retweets und Followern ausgewertet, um daraus abzuleiten, was die dominante öffentliche Meinung ist, welche Akteure die größte

politische Unterstützung genießen oder die besten Wahlchancen haben (für einen kritischen Überblick: Jungherr, 2016; Jungherr et al., 2017; Lewis & Molyneux, 2018, S. 16–17). Zumeist wurde in solchen Arbeiten nicht elaboriert, wie der Schluss von den Nutzungsdaten auf das interessierende Makrophänomen konzeptionell möglich ist (Jungherr et al., 2017, S. 339). Auch wurde bei solchen weitgehend unbegründeten Inferenzen der bereits länger bekannte Umstand der mangelnden Repräsentativität übergangen. Beispielsweise nutzt nur ein kleiner Teil der Bevölkerung Twitter, und davon ist wiederum nur ein kleiner Teil besonders aktiv (boyd & Crawford, 2012, S. 669; Driscoll & Walker, 2014, S. 1746; Hölzig, 2018; Jungherr, 2016).

Ähnliche Muster gibt es im Bereich der Forschung zu Rolle und Einfluss von Online-Nutzungsdaten im Journalismus. Die werberelevante Nutzungsmetrik der Page Views ist in vielen Redaktionen zu einem wichtigen Orientierungsmaßstab geworden (Belair-Gagnon, 2019; Fürst, 2018; Vu, 2014). Auf dieser Metrik bauen auch die Rankings der „meistgelesenen“ Artikel auf, die sich auf vielen Online-Nachrichtenseiten finden (Webster, 2014). Studien, die diese Rankings für eigene Datenanalysen verwenden, schließen von den Rankings der „meistgelesenen“ Artikel auf die „content preferences among the public“ (Boczkowski & Mitchellstein, 2013, S. 26) oder die „audience desires“ und „audience interests“ (Lee, Lewis, & Powers, 2014, S. 519–520; vgl. auch Wendelin, Engelmann, & Neubarth, 2014, S. 453; dazu kritisch Fürst, 2018, S. 191–195; Groot Kormelink & Costera Meijer, 2018, S. 670). Dabei wird nicht erläutert, inwiefern sich die Programmierung dieser Rankings gegebenenfalls unterscheidet (dazu instruktiv Zamith, 2018), wie mit aggregierten Zahlen der Nutzerzugriffe valide auf Interessen und Bedürfnisse geschlossen werden kann und inwiefern Rückschlüsse auf das Publikum als Einheit sinnvoll sind.

Ausgehend vom Einfluss dieser Daten auf den Journalismus wird zudem oft darauf geschlossen, dass sich das Publikum nun mit seinen Interessen stärker im Journalismus geltend machen könne (Stichwort „audience gatekeeping“ bzw. „changing power dynamics“, Lee et al., 2014, S. 520; Tandoc, 2019, S. 19; siehe auch Vu, 2014). Diese Inferenz wird durch vorliegende empirische Befunde und konzeptionelle Überlegungen deutlich infragegestellt (Costera Meijer, 2013; Fürst, 2018; Groot Kormelink & Costera Meijer, 2018; Hindman, 2018; von Krogh & Andersson, 2016; Webster, 2014, S. 87–88). Denn es gibt diverse Nutzungspraktiken, die bei der Messung von Page Views nicht erfasst werden. Der Zugriff auf einen Artikel bedeutet häufig auch nicht, dass dieser gelesen oder interessant gefunden wird – zumal NutzerInnen vor dem Zugriff kaum wissen können, ob ein Beitrag ihren Erwartungen entspricht. Insofern nimmt nicht „das Publikum“ mehr Einfluss auf den Journalismus, sondern eine rudimentäre, kommerziellen Zwecken dienende Nutzungsmetrik, die vorschnell als Abbild der Publikumsinteressen interpretiert wird.

D. Unkritische Übernahme von interessengeleiteten Superlativen und Generalisierungen

Während Rankings den Blick auf mehrere Phänomene richten und diesen jeweils einen „Rang“ zuweisen, rücken Superlative und Generalisierungen ein Phänomen in den Mittelpunkt und setzen es als absolut. Superlative haben Nachrichtenwert (Fürst, 2020; Fürst & Oehmer, 2021) und haben im Kontext des zunehmenden Wettbewerbs um Aufmerksamkeit (Webster, 2014) an Bedeutung gewonnen. Heintz (2019, S. 75) gibt dabei zu bedenken: „Damit eröffnet sich eine nicht enden wollende Spirale der Leistungssteigerung und Überbietung“. Vor diesem Hintergrund ist ein kritischer Blick auf die öffentliche Kommunikation von Technologie-, Daten-, Marktforschungs- und Medienunternehmen notwendig, die regelmäßig auf Superlative und Generalisierungen zurückgreifen. Zu fragen ist dabei stets, ob den Aussagen tatsächlich Daten zugrunde liegen, wie diese Daten gewonnen wurden und welche Akteure daran beteiligt waren.

Bisher werden diese Fragen in der Kommunikations- und Medienforschung oftmals übergangen. Einige Studien übernehmen die Narrative von Technologie- und Medienunternehmen, um die Relevanz des untersuchten Ereignisses zu unterstreichen. Beispielsweise werden mehrere TV-Debatten im Kontext der US Wahlen 2016 als „most tweeted day“ (Camaj & Northup, 2019, S. 34; siehe auch Robertson, Dutton, Ackland, & Peng, 2019, S. 105) oder Twitter „record“ (R. Berman, Melumad, Humphrey, & Meyer, 2019, S. 900; Jennings, Bramlett, McKinney, & Hardy, 2020, S. 1; Wasike, 2019, S. 252) bezeichnet. Die gleichen Superlative finden sich auch in Studien, die die Rolle von Twitter in der vorherigen Wahlperiode untersuchten (Coddington, Molyneux, & Lawrence, 2014, S. 392; Houston et al., 2013, S. 549; Molyneux, Mourão, & Coddington, 2017, S. 43; Shah et al., 2016, S. 1817). Auffällig daran ist: Die Aussagen werden stets mit Medienberichten oder PR-Aussagen von Twitter belegt. In allen Fällen werden keine oder nur wenige Daten beigebracht, deren Entstehung und Aussagekraft zudem unklar bleibt (z. B.: Wie wurden die Tweets identifiziert, die sich auf ein bestimmtes Ereignis beziehen?). Darüber hinaus variieren die Bezüge der behaupteten Superlative, so dass unklar bleibt, was jeweils die Vergleichsmaßstäbe und berücksichtigten Ereignisse für die Aussagen sind (siehe etwa Abb. 1, „political-event record“). Die mit den Superlativen signalisierte Außergewöhnlichkeit und Einzigartigkeit ist also weniger ein datenbasierter Fakt als eine rhetorische Strategie. Die Aussagen, auch jene aus den zitierten Medienberichten, gehen letztlich auf das Unternehmen Twitter zurück. Das Unternehmen postuliert mittels Superlativen eine kontinuierliche Bedeutungssteigerung seiner Plattform und ist dabei zugleich die einzige Instanz, die Wissen über die jeweils verglichenen Ereignisse sowie Zugang zu den gesamten Daten hat. Damit ist es für Außenstehende praktisch unmöglich, die Aussagen des Unternehmens zu prüfen – wohl aber kann man diese hinterfragen und als PR-Strategie kenntlich machen.

Ähnlich verhält es sich mit Generalisierungen wie „Facebook election“ und „Twitter election“. Diese werden in wissenschaftlichen Arbeiten übernommen, um die Relevanz von Studien zu begründen und „single platform studies“ (Bossetta, 2018, S. 474) zu legitimieren (siehe beispielsweise Houston et al., 2013, S. 548;

Johnson & Perlmutter, 2011; McKinney et al., 2014, S. 556, 560, 570; Zheng & Shahin, 2020, S. 338). Das Narrativ, dass eine Wahl bzw. der gesamte Wahlkampf wesentlich durch die Nutzung und den Einfluss einer spezifischen Plattform geprägt worden sei, wird von den Unternehmen selbst öffentlich kommuniziert und hat eine werbende Funktion (McKinney et al., 2014, S. 560; Molyneux et al., 2017, S. 43; Twitter, 2014, S. 8; 2016). Eine kritische Kommunikations- und Medienwissenschaft, die ihre gesellschaftliche Rolle wahrnimmt, wäre angesichts dessen insbesondere gefordert, solche Narrative zu hinterfragen und Unternehmensstrategien aufzudecken. Zudem ist es wichtig, dass die Forschung eigene Nutzungsdaten mit entsprechenden wissenschaftlichen Standards generiert und zitiert. In dem hier besprochenen Kontext sind dies etwa umfassende Befragungen, die aufzeigen, dass in 2016 nur knapp 10 Prozent der erwachsenen US-Bevölkerung mindestens einmal wöchentlich Twitter für Informationen zum Wahlkampf nutzte (Pew Research Center, 2016, S. 20) und knapp 6 Prozent der US-Wähler während der TV-Debatten auf Twitter aktiv waren (Holz, Akin, & Jamieson, 2016, S. 17).

In Medien und Wissenschaft kursierende Superlative der Mediennutzung sind nicht nur auf digitale Medien bezogen. So werden auch die Fernsehübertragungen der Präsidentschaftsdebatten beständig neu als Rekorde bezeichnet: In 2012 hätten so viele US-AmerikanerInnen eine Präsidentschaftsdebatte im Fernsehen verfolgt wie seit über drei Jahrzehnten nicht mehr (Shah et al., 2016, S. 1817); bei den Präsidentschaftsdebatten in 2016 seien es mehr ZuschauerInnen gewesen als jemals zuvor in der Geschichte (R. Berman et al., 2019, S. 900; Fuchs, 2018, S. 183, 254; Jennings et al., 2020, S. 1; Robertson et al., 2019, S. 105; Wasike, 2019, S. 252; Winneg & Jamieson, 2017, S. 362, 374, 376–377). Aus dieser unterstellten historischen Bedeutung wird eine besondere Relevanz der jeweils unternommenen Studie und der gewonnenen Erkenntnisse abgeleitet: „Because of the unprecedented numbers of viewers, the 2016 presidential debates can be considered [...] historic media events [...]. Therefore, the debates presented an opportunity to inform a vast number of voters about the candidates“ (Winneg & Jamieson, 2017, S. 374).

Diese postulierten Rekorde gehen auf öffentlichkeitswirksame Medienmitteilungen von Unternehmen der Markt- und Medienforschung zurück (Prior & Bougher, 2018) und wurden von diesen sogar im Vorfeld der Ereignisse als Erwartungen veröffentlicht (Nielsen Company, 2016, 15. August) und von zahlreichen Nachrichtenorganisationen verbreitet (Fürst, 2021; Serjeant, 2016, 22. September). Auch hier spielen also kommerzielle Interessen und PR-Strategien eine große Rolle. Das macht kritische Einordnungen dieser Nutzungsdaten notwendig. Schaut man sich beispielsweise die seit 1960 gemessenen Daten, so weit sie öffentlich sind, vergleichend an und berücksichtigt dabei die jeweilige Größe der US-Bevölkerung, so zeigt sich eine über die Jahrzehnte eher sinkende TV-Rezeption der Präsidentschaftsdebatten (Holz et al., 2016; Prior & Bougher, 2018, S. 830–831).

E. Verwendung von geschätzten Zahlen als gemessene Daten

Schließlich gibt es auch Zahlenmaterial, das lediglich auf Schätzungen basiert, aber von der Forschung so übernommen wird als handelte es sich um empirische Aussagen bzw. gemessene Daten. Am Beispiel der Forschung zur Diffusion des

Internets: Es liegen keine Daten zur weltweiten Anzahl der InternetnutzerInnen in den 1990er Jahren vor (Hargittai, 1999; Hoffman & Novak, 1994). Erste Studien zu einzelnen Ländern setzten Mitte der 1990er Jahre ein. Für die USA gibt es seit 1995 erste Daten, für Deutschland seit 1997. Viele andere Länder sowie international vergleichende Studien folgten erst später (als Überblick: Groebel & Gehrke, 2003). Jedoch kursierten Zahlen zur nationalen und globalen Internetnutzung bereits seit Mitte der 1990er Jahre in der Öffentlichkeit (Fürst, 2017). Zahlreiche Marktforschungs- und Beratungsunternehmen veröffentlichten hohe und rasant steigende Zahlen zur weltweiten Internetnutzung, fütterten damit euphorische Erwartungen an die Technologie und stimulierten Investitionen in die Branche (Jordan, 2001; Latzer, 1997, S. 41). Diese Zahlen wichen oftmals stark voneinander ab und können als „dubios“ (Latzer, 1997, S. 41) bezeichnet werden, denn zumeist ist nicht nachvollziehbar, wie sie entstanden sind (Batinic, 2000, S. 34). Teilweise basieren die veröffentlichten Zahlen darauf, dass die Anzahl der gemessenen Hosts mit einer bestimmten, nicht näher begründeten Zahl multipliziert wurde (Batinic, 2000, S. 34). „What happens when you multiply the number of hosts by the number of users per host? Depending on which numbers you select, you get something between 2.5 million and 32 million users.“ (Hoffman & Novak, 1994) Tatsächlich war die Anzahl der Nutzer pro Host sowie die Anzahl der genutzten Hosts pro Nutzer unbekannt (Hargittai, 1999; Hoffman & Novak, 1994). Auch konnte angenommen werden, dass sich die Anzahl der Nutzer pro Host von Land zu Land unterscheidet (Jordan, 2001).

Dennoch fanden Zahlen zur weltweiten Nutzung des Internets in den 1990er Jahren Eingang in die Kommunikations- und Medienforschung. Zum einen gibt es Arbeiten, die ohne Nennung einer Quelle und ohne weitere Erläuterungen solche Zahlen angeben, um die Relevanz ihres Untersuchungsgegenstands zu verdeutlichen (Bobbitt, 1995, S. 27), eine historische Entwicklung nachzuzeichnen (Faustich, 2012, S. 399; Neverla, 1998, S. 17) oder Evidenz für eine bestimmte Perspektive beizubringen: „With 25 million people estimated to be communicating on the Internet, should communication researchers now consider this network of networks a mass medium?“ (Morris & Ogan, 1996, S. 39). Zum anderen werden Zahlen zur Nutzung des Internets in den frühen und mittleren 1990er Jahren verwendet, um genaue Diffusionsverläufe aufzuzeigen. Wie die dabei verwendeten Zahlen entstanden sind, ist intransparent. Auch sind diese oft nicht (mehr) öffentlich zugänglich. Als Quelle für die Zahlen werden lediglich kurz die Namen oder Akronyme von Unternehmen und Organisationen genannt (wie bspw. NUA, NVA Ltd., International Telecommunication Union). WissenschaftlerInnen rechnen mit Zahlen von diesen Unternehmen und Organisationen oder erstellen auf deren Basis Graphiken – geben jedoch keine näheren Informationen zum dabei verwendeten Datensatz und zitierten Urheber (Andrés, Cuberes, Diouf, & Serebriskya, 2010; Marr, 2005, S. 65; Rogers, 2003, S. 347).

Zudem gibt es Zahlen, die sich zwischen Messungen und Schätzungen bewegen und weder für die Öffentlichkeit noch für die Wissenschaft direkt überprüfbar sind. Beispielsweise veröffentlichte das Marktforschungsunternehmen Media-Quant während des US-Wahlkampfs 2016 Zahlen zur Medienpräsenz von Trump im Vergleich zu anderen KandidatInnen („earned media“ der KandidatInnen an-

gegeben in US-Dollar) und beanspruchte dabei, dass diese Zahlen praktisch alle Medien umfassen (Print-, Rundfunk- und Onlineberichterstattung sowie Blogs und Social Media) und allgemeingültig sind. Diese Medienpräsenz berechnet sich aus der Thematisierung in verschiedensten Medien und der Anzahl der damit potenziell jeweils erreichten NutzerInnen.⁴ Die Zahlen wurden in zahlreichen Nachrichtenmedien zum Thema gemacht (im Überblick: <https://tinyurl.com/medi-aquant-news>). Allein der Bericht, der dazu im März 2016 in der *New York Times* erschien (Confessore & Yourish, 2016, 15. März), ist inzwischen in mehr als hundert wissenschaftlichen Arbeiten zitiert worden (siehe: <https://tinyurl.com/citing-NYT>) und dient der politischen Kommunikationsforschung als empirische Evidenz für die Aussage, dass Trump über alle Medien hinweg außerordentlich sichtbar war und die größte Reichweite erzielte (z. B.: Karpf, 2017, S. 199; Pickard, 2016, S. 118). Manche Forschungsarbeiten, die inzwischen bereits vielfach zitiert wurden, verwenden die Zahlen von MediaQuant sehr umfassend, auch in Form von tabellarischen und graphischen Darstellungen (Francia, 2018). Das Unternehmen MediaQuant legt jedoch weder offen, wie es Beiträge erfasst und codiert hat noch welche Medien konkret in die Analyse einbezogen wurden (Chadwick, 2017, S. 259). Die Zahlen und Aussagen von MediaQuant sind entsprechend nicht intersubjektiv überprüfbar. Indem die Forschung diese dennoch als empirische Belege zitiert, vertraut sie *blind* darauf, dass eine kommerziell ausgerichtete Organisation wissenschaftliche Methoden konsequent und systematisch anwendet. Zugleich vernachlässigt die Forschung, dass diese Zahlen eine starke Medienaufmerksamkeit erhielten und damit potenziell den Wahlkampf beeinflussten. Auch die wichtige Frage danach, wer diese Analyse in Auftrag gegeben hat und welchen Interessen sie dient, wurde bisher übergangen. Solche Fragen vermehrt zu adressieren, würde die gesellschaftskritische Rolle der Kommunikations- und Medienforschung stärken.

Eine ähnliche Problematik zeigt sich bei öffentlich kursierenden Nutzungszahlen zu globalen Medienereignissen, die im Fernsehen übertragen werden. Obwohl es keine Daten zum globalen Nutzungsverhalten gibt, sondern lediglich unterschiedlich angelegte nationale Messungen der Fernsehnutzung herangezogen und diskutiert werden können (Balnaves, O'Regan, & Goldsmith, 2011, S. 100–102; Bourdon & Méadel, 2014), greift die Kommunikations- und Medienwissenschaft auf globale Nutzungszahlen zurück, um konzeptionell grundlegende Aussa-

4 Im Bericht der *New York Times* (Confessore & Yourish, 2016, 15. März) heißt es: „The mentions are weighted by the reach of the media source, meaning how many people were likely to see it.“ Ein solches Verfahren ist hoch voraussetzungsreich, wenn zahlreiche Medientitel und unterschiedliche Medientypen einbezogen werden und müsste aus wissenschaftlicher Sicht genau dokumentiert werden. Gegenwärtig existiert die Webseite des Unternehmens MediaQuant jedoch nicht mehr. Teile dieser Webseite, die sich in den vergangenen Jahren auch mehrfach verändert hat, können noch über die *Wayback Machine* des Internet Archive eingesehen werden: http://web.archive.org/web/*https://www.mediaquant.net.

gen über ein Ereignis zu treffen⁵ sowie dessen besondere Relevanz aufzuzeigen (zum Beispiel: Dayan & Katz, 1992; Engell, 2008; Solves, Sánchez, & Rius, 2018). Für diese Zahlen werden häufig Medienberichte und PR-Mitteilungen der Veranstalter zitiert. Problematisch ist dies auch, weil die Größe des globalen Publikums bereits im Vorfeld des Ereignisses medial verbreitet und vordefiniert wird (Fürst, 2020; C. R. Martin & Reeves, 2001). Teilweise werden auch andere WissenschaftlerInnen zitiert, die wiederum die Nutzungszahlen mit Medien- und PR-Mitteilungen belegen oder gar keine Quelle dafür angeben (im Überblick: Fürst, 2020, S. 1526). Zahlreiche wissenschaftliche Arbeiten übernehmen auch die Zahlen aus den PR-Mitteilungen des Marktforschungsunternehmens Nielsen (zum Beispiel: <https://tinyurl.com/citing-Nielsen2008>). Einer PR-Mitteilung zufolge (Nielsen Company, 2008) hätten 4,7 Milliarden Menschen, also rund 70 Prozent der Weltbevölkerung, im Jahre 2008 die Olympischen Spiele im Fernsehen verfolgt. Wie diese Zahl konkret berechnet wurde, wird nicht offengelegt. In einer Fußnote ist jedoch vermerkt: „Estimates are based on data sourced across 37 markets from The Nielsen Company“ (ebd.). Die Reichweitenmessungen aus einigen (nicht angegebenen) Ländern wurden also verwendet, um eine Aussage über die weltweite Reichweite zu treffen. Es handelt sich bei dieser Zahl entsprechend um eine undurchschaubare Mischung aus gemessenen Daten und geschätzten Zahlen, die weder überprüfbar ist noch ein valides Vorgehen darstellt. Fälle, in denen Zahlen der kommerziellen Markt- und Medienforschung sich nicht als Daten, sondern als Schätzungen oder Fälschungen herausstellten, hat es bereits vor einigen Jahrzehnten gegeben (Balnaves et al., 2011, S. 74–77; Buzzard, 2015). Sie sind eine Mahnung daran, dass WissenschaftlerInnen die Nutzungsdaten, die von Unternehmen stammen, so weit wie möglich überprüfen müssen⁶ und in Verbindung mit weiteren Daten diskutieren sollten, deren Generierung nachvollziehbar ist und wissenschaftlichen Standards entspricht.

4. Fazit

Die Kommunikations- und Medienforschung sichert wissenschaftliche Qualitätsstandards durch den Einsatz und die Vermittlung wie Weiterentwicklung sozialwissenschaftlicher Methoden (z. B. Brosius et al., 2016; Matthes et al., 2011; Scheufele & Engelmann, 2009) und ethischer Selbstverpflichtungen (Filipović et al., 2016; Rakebrand & Schlütz, 2019). Zugleich findet jedoch ein zunehmender Rückgriff auf nicht-wissenschaftliche Nutzungsdaten statt. Dazu haben sich in verschiedenen Forschungsbereichen zu jeweils spezifischen Punkten kritische

5 Im Fall der einflussreichsten konzeptionellen Bestimmung von Medienereignissen („media events“) spielt die Größe des globalen Publikums eine wichtige Rolle. Dayan und Katz (1992, S. 13) definieren Medienereignisse als Ereignisse mit einem außergewöhnlich großen Publikum bzw. den „largest audiences in the history of the world“. Alle im Buch angegebenen globalen Nutzungszahlen zu Medienereignissen gehen jedoch auf Medienberichte zurück (dazu kritisch Fürst, 2020).

6 Diesen Überprüfungsmöglichkeiten sind häufig Grenzen gesetzt. Aussagekräftig ist es aber beispielsweise schon, wenn ein Sportverband oder ein Marktforschungsunternehmen die Berechnung von Nutzungszahlen auf Nachfrage nicht offenlegt und überrascht ist, dass danach gefragt wird (C. R. Martin & Reeves, 2001, S. 229–230).

Stimmen gemeldet. Eine breite und umfassende Diskussion ist dagegen bisher ausgeblieben. Auf Basis zahlreicher Literaturverweise und unter Einbeziehung verschiedener Bereiche der Kommunikations- und Medienforschung konnte ein erster Überblick gegeben und gezeigt werden, dass es mehrere bedeutende Problematiken im Umgang mit nicht-wissenschaftlichen Nutzungszahlen gibt. Oftmals ging es dabei um die quantitativ orientierte Forschung zu Social Media, politischer Kommunikation, Mediennutzung und Journalismus. Die analysierten Problematiken wurden aber auch in medienwissenschaftlichen, -historischen und -soziologischen Arbeiten festgestellt (siehe insb. Abschnitt E) und betreffen daher weite Teile des Fachs.

Die hier diskutierten Nutzungszahlen (A – E) sind mit starker Intransparenz verbunden, können häufig nicht überprüft werden und / oder müssten hinsichtlich ihrer Aussagekraft und der mit ihnen verfolgten kommerziellen Interessen und Strategien genau reflektiert werden. Wissenschaftliche Arbeiten greifen dennoch vielfach – und zumeist ohne kritische Reflexionen – auf solche Zahlen zurück, um die Relevanz des Untersuchungsgegenstands und der Forschungsfrage zu begründen, bestimmten Argumenten, konzeptionell grundlegenden Entscheidungen oder behandelten Entwicklungen und Perspektiven empirische Evidenz zu verleihen, die Relevanz der eigenen Ergebnisse zu verdeutlichen und weitreichende Inferenzen auf Makrophänomene zu ziehen. Zum Teil werden dabei wenige nicht-wissenschaftliche Nutzungsdaten nur kurz dargestellt, zum Teil werden viele Zahlen aufgegriffen oder graphisch dargestellt. Darüber hinaus gibt es auch Studien, die mit den nicht-wissenschaftlichen Nutzungsdaten rechnen oder Datenauswahl und Untersuchungsdesign an den Metriken und Rankings der Technologie- und Medienunternehmen ausrichten. Damit umfassen die identifizierten Problematiken praktisch alle Schritte der wissenschaftlichen Wissensproduktion.

Auffällig ist, dass mit dem Rückgriff auf die Nutzungsdaten häufig auch eine Übernahme der Narrative und rhetorischen Mittel verbunden ist, mit denen Unternehmen diese Daten kommunizieren. Die Kommunikations- und Medienforschung vernachlässigt bisher Fragen danach, welche Narrative von Unternehmen entwickelt und welche Strategien mit der Kommunikation von Nutzungszahlen verfolgt werden. Die mit den Zahlen und Narrativen intendierte Erzielung von öffentlicher Aufmerksamkeit und Medienberichterstattung wird eher selten in den Blick genommen (Fürst, 2021; McGregor, 2019; Prior & Bougher, 2018; van Dijk, 2013; van Dijk & Poell, 2013). Stattdessen zitieren viele ForscherInnen eben jene Medienberichte als Quelle für die verwendeten Zahlen, die mit der Veröffentlichung der Nutzungszahlen stimuliert werden sollten. Dies mag auch mit der kulturellen Bedeutung von Zahlen und ihrem Anschein von Objektivität zusammenhängen. Denn obschon die Gewinnung von Zahlen an eine Reihe von Entscheidungen und Voraussetzungen geknüpft ist, „werden sie in der Regel für die Sache selbst gehalten“ (Heintz, 2010, S. 170; siehe auch Lawson, 2021; Van Witsen, 2020, S. 1048).

Zwar sollten gerade SozialwissenschaftlerInnen den professionellen Reflex haben, Daten mit Blick auf die jeweilige Methodik zu reflektieren, mit der sie gewonnen wurden. Kommunikations- und MedienforscherInnen im Speziellen könnten zudem die Strukturen und Entwicklungen der Nachrichtenproduktion einbeziehen und reflektieren, dass journalistische Medien das Zahlenmaterial aus

PR-Mitteilungen oft unhinterfragt publizieren und zunehmend weniger Ressourcen für die Informationsüberprüfung haben (siehe Abschnitt 2). Doch die öffentliche Omnipräsenz, die bestimmte Metriken und Rankings gewonnen haben, sowie der gleichzeitig steigende wissenschaftliche Publikationsdruck und -rhythmus haben diesen Reflex offenbar geschwächt. Auch können sich bestimmte Umgangsweisen mit nicht-wissenschaftlichen Nutzungsdaten mit der Zeit als neuer Standard etablieren: Wer zum Beispiel bereits auf zehn bis zwanzig Studien zu „trending topics“ verweisen kann, spürt vermutlich weniger Notwendigkeit, die Berechnung, Bedeutung und Strategie dieser Rankings zu reflektieren und deren Verwendung zu rechtfertigen. Hinzu kommt, dass es in manchen, die Forschung interessierenden Bereichen noch keine wissenschaftlichen Daten gibt. So müssen sich Kommunikations- und MedienforscherInnen dann entscheiden, ob sie auf intransparente Zahlen zurückgreifen oder ob sie keine Daten für ein bestimmtes Argument, eine Perspektive oder die Relevanz eines Untersuchungsgegenstands vorbringen können. Wie im Fall der Forschung zur Diffusion des Internets oder zum Publikum von Medienereignissen deutlich wurde, sind die Anbieter nicht-wissenschaftlicher Zahlen immer schon schneller. Ihre Zahlen bzw. Schätzungen liegen vor ehe die ersten wissenschaftlichen Studien unternommen werden und prägen oft von Beginn an die wissenschaftlichen Annahmen und Perspektiven (siehe Abschnitt E).

Die vorliegende Analyse kann keine Aussage dazu machen, wie hoch der Anteil solcher Praktiken an der gesamten Forschung ist. Auch erhebt die Analyse keinen Anspruch, alle denkbaren Problematiken bei der Verwendung nicht-wissenschaftlicher Nutzungsdaten erfasst zu haben. Allerdings konnte mittels der fünf identifizierten Problematiken und an Hand von zahlreichen Studien und Beispielen ein Grundstein für eine umfassende Reflexion gelegt werden. Vermutlich gibt es viele WissenschaftlerInnen, die solche Daten bisher nicht direkt verwendet oder zitiert haben. Allerdings ist dies kein ausreichender Mechanismus für die Qualitätssicherung der Wissensproduktion. Denn jene Studien, bei denen problematische Praktiken festgestellt wurden, werden von weiteren Arbeiten zitiert und so zum Teil des tradierten Wissens. Verdeckt durch Zitationsketten werden sie am Ende unwissentlich auch von WissenschaftlerInnen tradiert, die nicht-wissenschaftliche Nutzungsdaten eigentlich kritisch einordnen würden.⁷

Die kritische Auseinandersetzung mit den eigenen Praktiken bietet der Kommunikations- und Medienforschung nicht nur die Möglichkeit, die wissenschaftliche Qualität dort zu sichern, wo sie gefährdet ist. Diese Auseinandersetzung kann darüber hinaus auch neue Fragen aufwerfen und gesellschaftskritische Perspektiven anregen. Es wurde ersichtlich, dass das Fach mit der Herausforderung konfrontiert ist, dass es einerseits so viele Nutzungsdaten gibt wie noch nie, diese aber andererseits vielfach intransparent und interessengeleitet sind. Die wissenschaftliche Forschung hat im Vergleich zu den großen Konzernen und Unternehmen einen deutlich eingeschränkteren Zugang zu Nutzungsdaten und kann mit diesen

7 Beispiele für solche Zitationsketten, durch die nicht-wissenschaftliche Nutzungsdaten ungeprüft in die Forschung eingehen, finden sich in Conway (2007, S. 40, 56), Fürst (2020, S. 1526) und Vancil und Pendell (1987, S. 18).

diesbezüglich nicht konkurrieren (Bruns, 2019; Driscoll & Walker, 2014; Hepp, 2016; Lohmeier, 2014; Schmitt, 2018). Sie kann aber eine wissenschaftlich wie gesellschaftlich relevante Rolle wahrnehmen, indem sie öffentlich kursierende Nutzungszahlen auf Plausibilität prüft, vor dem Hintergrund eigener Daten aus verschiedenen methodischen Zugängen reflektiert und die Interessen und Strategien der beteiligten Unternehmen und Organisationen offenlegt. Denn nicht-wissenschaftliche Nutzungszahlen werden offenkundig oftmals verbreitet, um Einfluss auf die Mediennutzung zu nehmen, Aufmerksamkeit zu generieren, bestimmte Publikumsvorstellungen zu evozieren und selbsterfüllende Prophezeiungen auszulösen (Fürst, 2017, 2020, 2021; Hartmann & Dohle, 2005; Jarren, 2015; Stauff & Thiele, 2007; Webster, 2014). Die Kommunikations- und Medienforschung kann in der Gesellschaft und im öffentlichen Diskurs zu mehr Reflexion darüber anregen, wie Menschen durch Zahlen repräsentiert und Öffentlichkeitsdynamiken (Fürst, 2021) mittels Zahlen stimuliert werden. Dies wäre ein Beitrag zur Stärkung des kritischen Reflexionspotenzials der Kommunikations- und Medienforschung, das in den jüngsten Jahren vermehrt eingefordert wird (Bakardjieva & Gehl, 2017; Dencik, 2020; Fuchs, 2021; Hepp, 2016; Jensen, 2019; Lewis & Molyneux, 2018; Livingstone, 2019).

Literatur

- Ahmad, M. I. (2016). The magical realism of body counts: How media credulity and flawed statistics sustain a controversial policy. *Journalism*, 17(1), 18–34. <https://doi.org/10.1177/1464884915593237>
- Andrés, L., Cuberes, D., Diouf, M., & Serebriskya, T. (2010). The diffusion of the Internet: A cross-country analysis. *Telecommunications Policy*, 34(5–6), 323–340. <https://doi.org/10.1016/j.telpol.2010.01.003>
- Anstead, N., & O’Loughlin, B. (2015). Social media analysis and public opinion: The 2010 UK general election. *Journal of Computer-Mediated Communication*, 20(2), 204–220. <https://doi.org/10.1111/jcc4.12102>
- Ausserhofer, J., & Maireder, A. (2013). National politics on Twitter: Structures and topics of a networked public sphere. *Information, Communication & Society*, 16(3), 291–314. <https://doi.org/10.1080/1369118X.2012.756050>
- Bakardjieva, M., & Gehl, R. W. (2017). Critical approaches to communication technology – The past five years. *Annals of the International Communication Association*, 41(3–4), 213–219. <https://doi.org/10.1080/23808985.2017.1374201>
- Baldwin-Philippi, J. (2019). Data-driven campaigning. In N. J. Stroud & S. C. McGregor (Hrsg.), *Digital discussions: How big data informs political communication* (S. 156–175). Routledge.
- Balnaves, M., O’Regan, T., & Goldsmith, B. (2011). *Rating the audience: The business of media*. London: Bloomsbury Academic. <https://doi.org/10.5040/9781849664622>
- Bastos, M. T., Raimundo, R. L. G., & Travitzki, R. (2013). Gatekeeping Twitter: Message diffusion in political hashtags. *Media, Culture & Society*, 35(2), 260–270. <https://doi.org/10.1177/0163443712467594>
- Batinic, B. (2000). Der Internet-Nutzer – ein rein theoretisches Konstrukt? In W. A. Mahle (Hrsg.), *Orientierung in der Informationsgesellschaft* (S. 33–44). UVK.

- Baur, N., & Blasius, J. (2019). Methoden der empirischen Sozialforschung – Ein Überblick. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 1–28). Springer VS. https://doi.org/10.1007/978-3-658-21308-4_1
- Beer, D. (2016a). How should we do the history of big data? *Big Data & Society*, 3(1), 1–10. <https://doi.org/10.1177/2053951716646135>
- Beer, D. (2016b). *Metric power*. Palgrave Macmillan. <https://doi.org/10.1057/978-1-137-55649-3>
- Beer, D. (2018). Envisioning the power of data analytics. *Information, Communication & Society*, 21(3), 465–479. <https://doi.org/10.1080/1369118X.2017.1289232>
- Belair-Gagnon, V. (2019). News on the fly: Journalist-audience online engagement success as a cultural matching process. *Media, Culture & Society*, 41(6), 757–773. <https://doi.org/10.1177/0163443718813473>
- Berman, E. P., & Hirschman, D. (2018). The sociology of quantification: Where are we now? *Contemporary Sociology: A Journal of Reviews*, 47(3), 2537–2266. <https://doi.org/10.1177/0094306118767649>
- Berman, R., Melumad, S., Humphrey, C., & Meyer, R. (2019). A tale of two Twitter-spheres: Political microblogging during and after the 2016 primary and presidential debates. *Journal of Marketing Research*, 56(6), 895–917. <https://doi.org/10.1177/0022243719861923>
- Bobbitt, R. (1995). An Internet primer for public relations. *Public Relations Quarterly*, 40(3), 27–32.
- Boczkowski, P. J., & Mitchelstein, E. (2013). *The news gap: When the information preferences of the media and the public diverge*. MIT Press.
- Bolin, G., & Velkova, J. (2020). Audience-metric continuity? Approaching the meaning of measurement in the digital everyday. *Media, Culture & Society*, 42(7–8), 1193–1209. <https://doi.org/10.1177/0163443720907017>
- Boone, G. M., Secci, J., & Gallant, L. M. (2018). Resistance: Active and creative political protest strategies. *American Behavioral Scientist*, 62(3), 353–374. <https://doi.org/10.1177/0002764217735623>
- Bossetta, M. (2018). The digital architectures of social media: Comparing political campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 U.S. Election. *Journalism & Mass Communication Quarterly*, 95(2), 471–496. <https://doi.org/10.1177/1077699018763307>
- Bourdon, J., & Méadel, C. (Hrsg.) (2014). *Television audiences across the world: Deconstructing the ratings machine*. Palgrave Macmillan. <https://doi.org/10.1057/9781137345103>
- boyd, d., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Brosius, H.-B., Haas, A., & Koschel, F. (2016). *Methoden der empirischen Kommunikationsforschung. Eine Einführung* (7., überarb. und akt. Aufl.). Springer VS. <https://doi.org/10.1007/978-3-531-19996-2>
- Bruns, A. (2019). After the ‘APIcalypse’: Social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, 22(11), 1544–1566. <https://doi.org/10.1080/1369118X.2019.1637447>

- Bruns, A. (2020). Big social data approaches in Internet studies: The case of Twitter. In J. Hunsinger, M. M. Allen, & L. Klasturp (Hrsg.), *Second international handbook of Internet research* (S. 65–81). Springer. https://doi.org/10.1007/978-94-024-1555-1_3
- Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, 20(1), 30–44. <https://doi.org/10.1080/1369118X.2016.1154086>
- Burk, V., & Grimmer, C. G. (2017). Die Fußball-WM 2014 im Fokus der Sozialen Medien – Information, Partizipation, Interaktion. In H. Ihle, M. Meyen, J. Mittag, & J.-U. Nie-land (Hrsg.), *Globales Mega-Event und nationaler Konflikt*. Die Fußball-WM 2014 in Medien und Politik (S. 83–105). Springer VS. https://doi.org/10.1007/978-3-658-16197-2_5
- Buzzard, K. S. F. (2015). The rise of market information regimes and the historical development of audience ratings. *Historical Journal of Film, Radio and Television*, 35(3), 511–517. <https://doi.org/10.1080/01439685.2015.1052219>
- Camaj, L., & Northup, T. (2019). Dual-screening the candidate image during presidential debates: The moderating role of Twitter and need to evaluate for the effects on candidate perceptions. *Journal of Broadcasting & Electronic Media*, 63(1), 20–38. <https://doi.org/10.1080/08838151.2019.1574117>
- Chadwick, A. (2017). *The hybrid media system: Politics and power* (2. Aufl.). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199759477.001.0001>
- Chadwick, A., & Stromer-Galley, J. (2016). Digital media, power, and democracy in parties and election campaigns: Party decline or party renewal? *The International Journal of Press/Politics*, 21(3), 283–293. <https://doi.org/10.1177/1940161216646731>
- Christin, A. (2020). What data can do: A typology of mechanisms. *International Journal of Communication*, 14, 1115–1134.
- Coddington, M., Molyneux, L., & Lawrence, R. G. (2014). Fact checking the campaign: How political reporters use Twitter to set the record straight (or not). *The International Journal of Press/Politics*, 19(4), 391–409. <https://doi.org/10.1177/1940161214540942>
- Confessore, N., & Yourish, K. (2016, 15. März). \$2 billion worth of free media for Donald Trump. *New York Times*. Abgerufen von <https://www.nytimes.com/2016/03/16/upshot/measuring-donald-trumps-mammoth-advantage-in-free-media.html>
- Conway, M. (2007). Before the bloggers: The upstart news technology of television at the 1948 political conventions. *American Journalism*, 24(1), 33–58. <https://doi.org/10.1080/08821127.2007.10678050>
- Costera Meijer, I. (2013). Valuable journalism: A search for quality from the vantage point of the user. *Journalism*, 14(6), 754–770. <https://doi.org/10.1177/1464884912455899>
- Curtin, P. A., & Maier, S. R. (2001). Numbers in the newsroom: A qualitative examination of a quantitative challenge. *Journalism & Mass Communication Quarterly*, 78(4), 720–738. <https://doi.org/10.1177/107769900107800407>
- Dayan, D., & Katz, E. (1992). *Media events: The live broadcasting of history*. Harvard University Press.
- Deller, R. (2011). Twittering on: Audience research and participation using Twitter. *Participations*, 8(1), 216–245.
- Dencik, L. (2020). Mobilizing media studies in an age of datafication. *Television & New Media*, 21(6), 568–573. <https://doi.org/10.1177/1527476420918848>

- Diaz-Bone, R. (2018). Neue Ausschlussdynamiken durch Big Data-generierte Unsichtbarkeiten, Inkohärenzen und ungleiche Zeitlichkeiten. In D. Houben & B. Priel (Hrsg.), *Datengesellschaft. Einsichten in die Datafizierung des Sozialen* (S. 207–229). Transcript. <https://doi.org/10.14361/9783839439579-009>
- Driscoll, K., & Walker, S. (2014). Working within a black box: Transparency in the collection and production of big Twitter data. *International Journal of Communication*, 8, 1745–1764.
- Drori, G. S. (2020). Hasn't institutional theory always been critical?! *Organization Theory*, 1(1), 1–9. <https://doi.org/10.1177/2631787719887982>
- Engell, L. (2008). Das Mondprogramm. Wie das Fernsehen das größte Ereignis aller Zeiten erzeugte. In F. Lenger & A. Nünning (Hrsg.), *Medienereignisse der Moderne* (S. 150–171). Wissenschaftliche Buchgesellschaft.
- Faulstich, W. (2012). *Die Mediengeschichte des 20. Jahrhunderts*. München: Wilhelm Fink.
- Filipović, A., Klaus, E., & Strippel, C. (2016). Der Ethik-Kodex der DGPK. *Medien & Kommunikationswissenschaft (M&K)*, 64(1), 93–108. <https://doi.org/10.5771/1615-634X-2016-1-93>
- Florini, S. (2014). Tweets, tweeps, and signifyin': Communication and cultural performance on "Black Twitter". *Television & New Media*, 15(3), 223–237. <https://doi.org/10.1177/1527476413480247>
- Francia, P. L. (2018). Free media and Twitter in the 2016 presidential election: The unconventional campaign of Donald Trump. *Social Science Computer Review*, 36(4), 440–455. <https://doi.org/10.1177/0894439317730302>
- Freelon, D. (2014). On the interpretation of digital trace data in communication and social computing research. *Journal of Broadcasting & Electronic Media*, 58(1), 59–75. <https://doi.org/10.1080/08838151.2013.875018>
- Freelon, D., & Karpf, D. (2015). Of big birds and bayonets: Hybrid Twitter interactivity in the 2012 Presidential debates. *Information, Communication & Society*, 18(4), 390–406. <https://doi.org/10.1080/1369118X.2014.952659>
- Fuchs, C. (2018). *Digital demagogue: Authoritarian capitalism in the age of Trump and Twitter*. Pluto Press.
- Fuchs, C. (2021). *Soziale Medien und Kritische Theorie. Eine Einführung* (2., vollst. überarb. Aufl.). UVK.
- Fürst, S. (2014). Werbeperspektiven in der Krise der Printmedien – Zur Inszenierung von Verbreitung und Wirkungsmacht der Presse. In H. Schramm & J. Knoll (Hrsg.), *Innovation der Persuasion. Die Qualität der Werbe- und Markenkommunikation in neuen Medienwelten* (S. 276–294). Herbert von Halem.
- Fürst, S. (2017). Die Etablierung des Internets als Self-Fulfilling Prophecy? Zur Rolle der öffentlichen Kommunikation bei der Diffusion neuer Medien. *Medien & Zeit*, 32(2), 43–55.
- Fürst, S. (2018). Popularität statt Relevanz? Die journalistische Orientierung an Online-Nutzungsdaten. In T. Mämecke, J.-H. Passoth & J. Wehner (Hrsg.), *Bedeutende Daten. Modelle, Verfahren und Praxis der Vermessung und Verdattung im Netz* (S. 171–204). Springer VS. https://doi.org/10.1007/978-3-658-11781-8_9
- Fürst, S. (2020). "The whole world watching"? How news media create the myth of an audience of billions and foster imagined communities. *International Journal of Communication*, 14, 1524–1541.

- Fürst, S. (2021). Neue Öffentlichkeitsdynamiken: Zu selbstverstärkenden, plattformübergreifenden Effekten von ‚Popularität‘. In M. Eisenegger, M. Prinzing, P. Ettinger, & R. Blum (Hrsg.), *Digitaler Strukturwandel der Öffentlichkeit. Historische Verortung, Modelle und Konsequenzen* (S. 339–359). Springer VS. https://doi.org/10.1007/978-3-658-32133-8_19
- Fürst, S., Meißner, M., Hofstetter, B., Puppis, M., & Schönhagen, P. (2017). Gefährdete Autonomie? Kontinuität und Wandel der journalistischen Berichterstattungsfreiheit und redaktioneller Arbeitsbedingungen in der Schweiz. In I. Stapf, M. Prinzing, & A. Filipović (Hrsg.), *Gesellschaft ohne Diskurs? Digitaler Wandel und Journalismus aus medienethischer Perspektive* (S. 219–236). Nomos. <https://doi.org/10.5771/9783845279824-219>
- Fürst, S., & Oehmer, F. (2021). Attention for attention hotspots: Exploring the newsworthiness of public response in the metric society. *Journalism Studies*, 22(6), 799–819. <https://doi.org/10.1080/1461670X.2021.1889396>
- Fürst, S., & Schönhagen, P. (2020). The “mediated social communication” approach: An early discursive mass communication model. In M. Bergman, K. Kirtiklis, & J. Siebers (Hrsg.), *Models of communication: Theoretical and philosophical approaches* (S. 113–132). Routledge. <https://doi.org/10.4324/9781315231402-8>
- Gainous, J., & Wagner, K. M. (2014). *Tweeting to power: The social media revolution in American politics*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199965076.001.0001>
- Gerlitz, C., & Rieder, B. (2018). Tweets are not created equal: Investigating Twitter’s client ecosystem. *International Journal of Communication*, 12, 528–547.
- Gil de Zúñiga, H., Garcia-Perdomo, V., & McGregor, S. C. (2015). What is second screening? Exploring motivations of second screen use and its effect on online political participation. *Journal of Communication*, 65(5), 793–815. <https://doi.org/10.1111/jcom.12174>
- Gillespie, T. (2011). Can an algorithm be wrong? Twitter Trends, the specter of censorship, and our faith in the algorithms around us. *Culture Digitally*. Abgerufen von <http://culturedigitally.org/2011/10/can-an-algorithm-be-wrong/>
- Gillespie, T. (2016). #Trendingistrending: When algorithms become culture. In R. Seyfert & J. Roberge (Hrsg.), *Algorithmic cultures: Essays on meaning, performance and new technologies* (S. 52–75). Routledge.
- Glerup, C., & Horst, M. (2014). Mapping ‘social responsibility’ in science. *Journal of Responsible Innovation*, 1(1), 31–50. <https://doi.org/10.1080/23299460.2014.882077>
- Gorwa, R., & Guilbeault, D. (2020). Unpacking the social media bot: A typology to guide research and policy. *Policy & Internet*, 12(2), 225–248. <https://doi.org/10.1002/poi3.184>
- Gottfried, J. A., Hardy, B. W., Holbert, R. L., Winneg, K. M., & Jamieson, K. H. (2017). The changing nature of political debate consumption: Social media, multitasking, and knowledge acquisition. *Political Communication* 34(2), 172–199. <https://doi.org/10.1080/10584609.2016.1154120>
- Grittmann, E., & Drücke, R. (2016). Wissenschaft als Gesellschaftskritik – ein Einwurf. *Medien & Kommunikationswissenschaft (M&K)*, 64(1), 109–113. <https://doi.org/10.5771/1615-634X-2016-1-109>

- Groebe, J., & Gehrke, G. (2003). *Internet 2002: Deutschland und die digitale Welt*. Leske + Budrich. <https://doi.org/10.1007/978-3-322-93456-7>
- Groot Kormelink, T., & Costera Meijer, I. (2018). What clicks actually mean: Exploring digital news user practices. *Journalism*, 19(5), 668–683. <https://doi.org/10.1177/1464884916688290>
- Häder, M. (2019). *Empirische Sozialforschung. Eine Einführung* (4. Aufl.). Springer VS. <https://doi.org/10.1007/978-3-658-26986-9>
- Haim, M., Kümpel, A. S., & Brosius, H.-B. (2018). Popularity cues in online media: A review of conceptualizations, operationalizations, and general effects. *Studies in Communication and Media (SCM)*, 7(2), 188–207. <https://doi.org/10.5771/2192-4007-2018-2-58>
- Hargittai, E. (1999). Weaving the western web: Explaining differences in Internet connectivity among OECD countries. *Telecommunications Policy*, 23, 701–718. [https://doi.org/10.1016/S0308-5961\(99\)00050-6](https://doi.org/10.1016/S0308-5961(99)00050-6)
- Harrison, S. (2021). Journalism and numeracy in context: Four case studies. *Numeracy: Advancing Education in Quantitative Literacy*, 14(2), 1–16. <https://doi.org/10.5038/1936-4660.14.2.1371>
- Hartmann, T., & Dohle, M. (2005). Publikumsvorstellungen im Rezeptionsprozess. *Publizistik*, 50(3), 287–303. <https://doi.org/10.1007/s11616-005-0133-5>
- Heintz, B. (2010). Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs. *Zeitschrift für Soziologie*, 39(3), 162–181. <https://doi.org/10.1515/zfsoz-2010-0301>
- Heintz, B. (2019). Vom Komparativ zum Superlativ. Eine kleine Soziologie der Rangliste. In S. Nicolae, M. Endreß, O. Berli, & D. Bischur (Hrsg.), *(Be)Werten. Beiträge zur sozialen Konstruktion von Wertigkeit* (S. 45–79). Springer VS. https://doi.org/10.1007/978-3-658-21763-1_3
- Hepp, A. (2016). Kommunikations- und Medienwissenschaft in datengetriebenen Zeiten. *Publizistik*, 61(3), 225–246. <https://doi.org/10.1007/s11616-016-0263-y>
- Hepp, A., Breiter, A., & Friemel, T. N. (2018). Digital traces in context. *International Journal of Communication*, 12, 439–449.
- Hindman, M. (2018). *The Internet trap: How the digital economy builds monopolies and undermines democracy*. Princeton University Press.
- Hoffman, D. L., & Novak, T. P. (1994). Wanted: Net.Census. *Wired*. Abgerufen von <http://www.wired.com/1994/11/hoffman-if/>
- Hölig, S. (2018). Eine meinungsstarke Minderheit als Stimmungsbarometer?! Über die Persönlichkeitseigenschaften aktiver Twitterer. *Medien & Kommunikationswissenschaft (M&K)*, 66(2), 140–169. <https://doi.org/10.5771/1615-634X-2018-2-140>
- Holz, J., Akin, H., & Jamieson, K. H. (2016). Presidential debates: What's behind the numbers? *Annenberg Public Policy Center*. Abgerufen von <https://www.annenbergpublicpolicycenter.org/publication/presidential-debates-whats-behind-the-numbers/>
- Hopke, J. E. (2015). Hashtagging politics: Transnational anti-fracking movement Twitter practices. *Social Media + Society*, 1(2), 1–12. <https://doi.org/10.1177/20563305115605521>
- Houston, J. B., McKinney, M. S., Hawthorne, J., & Spialek, M. L. (2013). Frequency of tweeting during presidential debates: Effect on debate attitudes and knowledge. *Communication Studies*, 64(5), 548–560. <https://doi.org/10.1080/10510974.2013.832693>

- Ignatow, G. (2020). *Sociological theory in the digital age*. Routledge. <https://doi.org/10.4324/9780429292804>
- Iliadis, A., & Russo, F. (2016). Critical data studies: An introduction. *Big Data & Society*, 3(2), 1–7. <https://doi.org/10.1177/2053951716674238>
- Jarren, O. (2015). Ordnung durch Medien? In K.-D. Altmeppen, P. Donges, M. Künzler, M. Puppis, U. Röttger, & H. Wessler (Hrsg.), *Soziale Ordnung durch Kommunikation?* (S. 29–50). Nomos. <https://doi.org/10.5771/9783845248035-27>
- Jarren, O., Künzler, M., & Puppis, M. (2019). Medienstrukturen als zentrales Forschungsfeld der Kommunikationswissenschaft. *Publizistik*, 64(4), 417–426. <https://doi.org/10.1007/s11616-019-00517-y>
- Jennings, F. J., Bramlett, J. C., McKinney, M. S., & Hardy, M. M. (2020). Tweeting along partisan lines: Identity-motivated elaboration and presidential debates. *Social Media + Society*, 6(4), 1–12. <https://doi.org/10.1177/2056305120965518>
- Jensen, K. B. (2019). The double hermeneutics of audience research. *Television & New Media*, 20(2), 142–154. <https://doi.org/10.1177/1527476418811103>
- Johnson, T. J., & Perlmutter, D. D. (Hrsg.) (2011). *New media, campaigning and the 2008 Facebook election*. Routledge. <https://doi.org/10.4324/9781315872650>
- Jordan, T. (2001). Measuring the Internet: Host counts versus business plans. *Information, Communication & Society*, 4(1), 34–53. <https://doi.org/10.1080/13691180122452>
- Jungherr, A. (2016). Twitter use in election campaigns: A systematic literature review. *Journal of Information Technology & Politics*, 13(1), 72–91. <https://doi.org/10.1080/19331681.2015.1132401>
- Jungherr, A., Schoen, H., Posegga, O., & Jürgens, P. (2017). Digital trace data in the study of public opinion: An indicator of attention toward politics rather than political support. *Social Science Computer Review*, 35(3), 336–356. <https://doi.org/10.1177/0894439316631043>
- Karpf, D. (2012). Social science research methods in Internet time. *Information, Communication & Society*, 15(5), 639–661. <https://doi.org/10.1080/1369118X.2012.665468>
- Karpf, D. (2017). Digital politics after Trump. *Annals of the International Communication Association*, 41(2), 198–207. <https://doi.org/10.1080/23808985.2017.1316675>
- Keller, T. R., & Klinger, U. (2021). The needle in a haystack: Finding social bots on Twitter. In E. Hargittai (Hrsg.), *How empirical social science gets done in the digital age* (S. 30–49). Columbia University Press. <https://doi.org/10.7312/harg18876-003>
- Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society*, 20(1), 14–29. <https://doi.org/10.1080/1369118X.2016.1154087>
- Kreiss, D., & McGregor, S. C. (2018). Technology firms shape political communication: The work of Microsoft, Facebook, Twitter, and Google with campaigns during the 2016 U.S. presidential cycle. *Political Communication*, 35(2), 155–177. <https://doi.org/10.1080/10584609.2017.1364814>
- Krüger, U., & Seignani, S. (2019). Einleitung. In U. Krüger & S. Seignani (Hrsg.), *Ideologie, Kritik, Öffentlichkeit: Verhandlungen des Netzwerks Kritische Kommunikationswissenschaft* (S. 1–10). Universität Leipzig.
- Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? Abgerufen von <http://dl.acm.org/citation.cfm?id=1772751>

- Latzer, M. (1997). *Mediamatik – Die Konvergenz von Telekommunikation, Computer und Rundfunk*. Westdeutscher Verlag. <https://doi.org/10.1007/978-3-322-86878-7>
- Lawson, B. T. (2021). Hiding behind databases, institutions and actors: How journalists use statistics in reporting humanitarian crises. *Journalism Practice*. Advance online publication. <https://doi.org/10.1080/17512786.2021.1930106>
- Lee, A. M., Lewis, S. C., & Powers, M. (2014). Audience clicks and news placement: A study of time-lagged influence in online journalism. *Communication Research*, 41(4), 505–530. <https://doi.org/10.1177/0093650212467031>
- Lewis, S. C., & Molyneux, L. (2018). A decade of research on social media and journalism: Assumptions, blind spots, and a way forward. *Media and Communication*, 6(4), 11–23. <http://doi.org/10.17645/mac.v6i4.1562>
- Livingstone, S. (2019). Audiences in an age of datafication: Critical questions for media research. *Television & New Media*, 20(2), 170–183. <https://doi.org/10.1177/1527476418811118>
- Löblich, M. (2010). Die empirisch-sozialwissenschaftliche Wende. Ein Beitrag zur historischen und kognitiven Identität der Kommunikationswissenschaft. *Medien & Kommunikationswissenschaft (M&K)*, 58(4), 544–562. <https://doi.org/10.5771/1615-634x-2010-4-544>
- Lohmeier, C. (2014). The researcher and the never-ending field: Reconsidering big data and digital ethnography. In M. Hand & S. Hillyard (Hrsg.), *Big data? Qualitative approaches to digital research* (S. 75–89). Emerald Group. <https://doi.org/10.1108/S1042-319220140000013005>
- Lomborg, S. (2017). A state of flux: Histories of social media research. *European Journal of Communication*, 32(1), 6–15. <https://doi.org/10.1177/0267323116682807>
- Luhmann, N. (1997). *Die Gesellschaft der Gesellschaft*. Suhrkamp.
- Mahrt, M., & Scharrow, M. (2013). The value of big data in digital media research. *Journal of Broadcasting & Electronic Media*, 57(1), 20–33. <https://doi.org/10.1080/08838151.2012.761700>
- Marchetti, R., & Ceccobelli, D. (2016). Twitter and television in a hybrid media system: The 2013 Italian election campaign. *Journalism Practice*, 10(5), 626–644. <https://doi.org/10.1080/17512786.2015.1040051>
- Marr, M. (2005). *Internetzugang und politische Informiertheit. Zur digitalen Spaltung der Gesellschaft*. UVK.
- Marres, N. (2017). *Digital sociology: The reinvention of social research*. Polity.
- Martin, C. R., & Reeves, J. L. (2001). The whole world isn't watching (but we thought they were): The Super Bowl and United States solipsism. *Sport in Society*, 4(2), 213–236. <https://doi.org/10.1080/713999822>
- Martin, F., & Dwyer, T. (2019). *Sharing news online: Commendary cultures and social media news ecologies*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-17906-9>
- Matthes, J., Kuhlmann, C., Gehrau, V., Jandura, O., Möhring, W., Vogelgesang, J., & Wunsch, C. (2011). Zur Methodenausbildung in kommunikationswissenschaftlichen Bachelor- und Masterstudiengängen. Empfehlungen einer Kommission im Auftrag der Fachgruppe Methoden der Deutschen Gesellschaft für Publistik- und Kommunikationswissenschaft. *Publizistik*, 56(4), 461–481. <https://doi.org/10.1007/s11616-011-0133-6>
- Mau, S. (2019). *The metric society: On the quantification of the social*. Cambridge: Polity.

- McGregor, S. C. (2019). Social media as public opinion: How journalists use social media to represent public opinion. *Journalism*, 20(8), 1070–1086. <https://doi.org/10.1177/1464884919845458>
- McKinney, M. S., Houston, J. B., & Hawthorne, J. (2014). Social watching a 2012 Republican presidential primary debate. *American Behavioral Scientist*, 58(4), 556–573. <https://doi.org/10.1177/0002764213506211>
- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.
- Molyneux, L., Mourão, R. R., & Coddington, M. (2017). US political journalists' use of Twitter: Lessons from 2012 and a look ahead. In R. Davis, C. Holtz-Bacha, & M. R. Just (Hrsg.), *Twitter and elections around the world: Campaigning in 140 characters or less* (S. 43–56). Routledge. <https://doi.org/10.4324/9781315669113-14>
- Morris, M., & Ogan, C. (1996). The Internet as mass medium. *Journal of Communication*, 46(1), 39–50. <https://doi.org/10.1111/j.1460-2466.1996.tb01460.x>
- Muhle, F. (2020). Alles Bots? Ein Vorschlag zur Typisierung (teil-)automatisierter politischer Accounts auf Twitter. In S. Breidenbach, P. Klimczak, & C. Petersen (Hrsg.), *Soziale Medien. Interdisziplinäre Zugänge zur Onlinekommunikation* (S. 45–70). Springer Vieweg. https://doi.org/10.1007/978-3-658-30702-8_3
- Muhle, F., Ackland, R., & Graham, T. (2018). Socialbots in politischen Online-Konversationen. Eine (überschätzte) Gefahr für demokratische Willensbildung im Internet? *Zeitschrift für Parlamentsfragen*, 49, 618–638. <https://doi.org/10.5771/0340-1758-2018-3-618>
- Neverla, I. (1998). Das Medium denken. Zur sozialen Konstruktion des Netz-Mediums. In I. Neverla (Hrsg.), *Das Netz-Medium. Kommunikationswissenschaftliche Aspekte eines Mediums in Entwicklung* (S. 17–35). Springer. https://doi.org/10.1007/978-3-663-07775-6_2
- Nguyen, A. (2018). Exciting times in the shadow of the 'post-truth' era: News, numbers and public opinion in a data-driven world. In A. Nguyen (Hrsg.), *News, numbers and public opinion in a data-driven world* (S. 1–15). Bloomsbury. <https://doi.org/10.5040/9781501330384.0005>
- Nielsen Company. (2008). The final tally: 4.7 billion tunes in to Beijing 2008 – More than two in three people worldwide. *Nielsen*. Abgerufen von <https://tinyurl.com/the-final-tally>
- Nielsen Company. (2016, 15. August). From your house to the White House: The top presidential debates. Abgerufen von <https://tinyurl.com/Nielsen-debates>
- Opp, K.-D. (2014). *Methodologie der Sozialwissenschaften. Einführung in Probleme ihrer Theorienbildung und praktischen Anwendung*. Springer VS. <https://doi.org/10.1007/978-3-658-01911-2>
- Pedro-Carañana, J., Alvarado-Vivas, S., & López-López, J. S. (2020). Agenda-setting and power relations during the 2018 Colombian election campaign on Twitter. *The Journal of International Communication*, 26(2), 260–280. <https://doi.org/10.1080/13216597.2020.1806900>
- Peifer, J. T., & Holbert, R. L. (2013). Developing a systematic assessment of humor in the context of the 2012 U.S. general election debates. *Argumentation and Advocacy*, 49(4), 286–300. <https://doi.org/10.1080/00028533.2013.11821803>

- Pew Research Center. (2016). *The 2016 presidential campaign – A news event that's hard to miss*. Abgerufen von <http://tinyurl.com/campaign-hard-to-miss>
- Pickard, V. (2016). Media failures in the age of Trump. *The Political Economy of Communication*, 4(2), 118–122.
- Pickard, V. (2020). *Democracy without journalism? Confronting the misinformation society*. Oxford University Press. <https://doi.org/10.1093/oso/9780190946753.001.0001>
- Popper, K. R. (1984). *Logik der Forschung* (8., weiter verbesserte u. vermehrte Aufl.). Mohr.
- Popper, K. R. (2017 [1962]). Die Logik der Sozialwissenschaften. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69(supplement issue 1), 215–228. <https://doi.org/10.1007/s11577-017-0425-6>
- Porter, T. M. (1996). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton: Princeton University Press.
- Prietzl, B., & Houben, D. (2018). Einführung. Soziologische Perspektiven auf die Datafizierung der Gesellschaft. In D. Houben & B. Prietzl (Hrsg.), *Datengesellschaft. Einsichten in die Datafizierung des Sozialen* (S. 7–32). Transcript. <https://doi.org/10.14361/9783839439579-001>
- Prim, R., & Tilmann, H. (1997). *Grundlagen einer kritisch-rationalen Sozialwissenschaft: Studienbuch zur Wissenschaftstheorie Karl R. Poppers* (7., erw. u. überarb. Aufl.). Quelle & Meyer.
- Prior, M., & Bougher, L. D. (2018). “Like they’ve never, ever seen in this country”? Political interest and voter engagement in 2016. *Public Opinion Quarterly*, 82(S1), 822–842. <https://doi.org/10.1093/poq/nfy002>
- Probst, C., Buhmann, A., Ingenhoff, D., & Lepori, B. (2019). Evolution of a field: Swiss media and communication studies. *Studies in Communication Sciences (SComS)*, 19(1), 7–23. <https://doi.org/10.24434/j.scoms.2019.01.002>
- Proferes, N. (2017). Information flow solipsism in an exploratory study of beliefs about Twitter. *Social Media + Society*, 3(1), 1–17. <https://doi.org/10.1177/2056305117698493>
- Proferes, N., & Summers, E. (2019). Algorithms and agenda-setting in Wikileaks’ #Podestamails release. *Information, Communication & Society*, 22(11), 1630–1645. <https://doi.org/10.1080/1369118X.2019.1626469>
- Rakebrand, T., & Schlütz, D. (2019). Forschungsethik in der kommunikations- und medienwissenschaftlichen Hochschullehre. Erkenntnisse aus der Praxis. *Medien & Kommunikationswissenschaft (M&K)*, 67(2), 187–205. <https://doi.org/10.5771/1615-634X-2019-2-187>
- Robertson, C. T., Dutton, W. H., Ackland, R., & Peng, T.-Q. (2019). The democratic role of social media in political debates: The use of Twitter in the first televised US presidential debate of 2016. *Journal of Information Technology & Politics*, 16(2), 105–118. <https://doi.org/10.1080/19331681.2019.1590283>
- Roeh, I., & Feldman, S. (1984). The rhetoric of numbers in front-page journalism: How numbers contribute to the melodramatic in the popular press. *Text & Talk*, 4(4), 347–368. <https://doi.org/10.1515/text.1.1984.4.4.347>
- Rogers, E. M. (2003). *Diffusion of innovations* (5. Aufl.). Free Press.
- Santini, R. M., Salles, D., Tucci, G., Ferreira, F., & Graef, F. (2020). Making up audience: Media bots and the falsification of the public sphere. *Communication Studies*, 71(3), 466–487. <https://doi.org/10.1080/10510974.2020.1735466>

- Scheufele, B., & Engelmann, I. (2009). *Empirische Kommunikationsforschung*. UVK.
- Schill, D., & Kirk, R. (2015). Issue debates in 140 characters: Online talk surrounding the 2012 debates. In J. A. Hendricks (Hrsg.), *Presidential campaigning and social media: An analysis of the 2012 campaign* (S. 198–217). Oxford University Press.
- Schmitt, M. (2018). Die Soziologie in Zeiten von Big Data. Angebote der Relationalen Soziologie. In D. Houben & B. Prietl (Hrsg.), *Datengesellschaft. Einsichten In die Datafizierung des Sozialen* (S. 299–319). Transcript. <https://doi.org/10.14361/9783839439579-013>
- Schönhagen, P., & Trebbe, J. (2005). Der Stellenwert eines sozialwissenschaftlichen Methodenkanons in der Kommunikations- und Medienwissenschaft der Schweiz. *Medienwissenschaft Schweiz*, 2/2005, 42–48.
- Schroeder, A. (2016). *Presidential debates: Risky business on the campaign trail* (3. Aufl.). Columbia University Press. <http://doi.org/10.7312/columbia/9780231170574.001.0001>
- Serjeant, J. (2016, 22. September). Presidential debate could score Super Bowl-sized audience. *Reuters*. Abgerufen von <https://tinyurl.com/yds7snwd>
- Shah, D. V., Hanna, A., Bucy, E. P., Lassen, D. S., Van Thomme, J., Bialik, K., . . . Pevehouse, J. C. W. (2016). Dual screening during presidential debates: Political nonverbals and the volume and valence of online expression. *American Behavioral Scientist*, 60(14), 1816–1843. <https://doi.org/10.1177/0002764216676245>
- Sharp, A. (2012). Dispatch from the Denver debate. *Twitter Blog*. Abgerufen von https://blog.twitter.com/en_us/a/2012/dispatch-from-the-denver-debate.html
- Solves, J., Sánchez, S., & Rius, I. (2018). The prince and the pauper: Journalistic culture and Paralympic Games in the Spanish print press. *Journalism*, 19(12), 1713–1728. <https://doi.org/10.1177/1464884916671894>
- Splichal, S. (2008). Why be critical? *Communication, Culture & Critique*, 1(1), 20–30. <https://doi.org/10.1111/j.1753-9137.2007.00003.x>
- Stark, B. (2013). Fragmentierung Revisited: eine theoretische und methodische Evaluation im Internetzeitalter. In W. Seufert & F. Sattelberger (Hrsg.), *Langfristiger Wandel von Medienstrukturen. Theorie, Methoden, Befunde* (S. 199–218). Nomos. <https://doi.org/10.5771/9783845249278-199>
- Stauff, M., & Thiele, M. (2007). Mediale Infografiken. Zur Popularisierung der Verdattung von Medien und ihrem Publikum. In I. Schneider & I. Otto (Hrsg.), *Formationen der Mediennutzung II: Strategien der Verdattung* (S. 251–267). Transcript. <https://doi.org/10.14361/9783839407424-014>
- Steininger, C., & Hummel, R. (2015). *Wissenschaftstheorie der Kommunikationswissenschaft*. de Gruyter. <https://doi.org/10.1515/9783486719109>
- Tandoc, E. C. (2019). *Analyzing analytics: Disrupting journalism one click at a time*. Routledge. <https://doi.org/10.4324/9781138496538>
- Taubert, N., & Weingart, P. (2016). Wandel des wissenschaftlichen Publizierens – eine Heuristik zur Analyse rezenter Wandlungsprozesse. In P. Weingart & N. Taubert (Hrsg.), *Zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung* (S. 3–38). De Gruyter. <https://doi.org/10.1515/9783110448115-001>
- Tausendpfund, M. (2018). *Quantitative Methoden in der Politikwissenschaft. Eine Einführung*. Springer VS. <https://doi.org/10.1007/978-3-658-20698-7>

- Turow, J., & Couldry, N. (2018). Media as data extraction: Towards a new map of a transformed communications field. *Journal of Communication*, 68(2), 415–423. <https://doi.org/10.1093/joc/jqx011>
- Twitter. (2010). *To trend or not to trend*. Twitter Blog. Abgerufen von https://blog.twitter.com/en_us/a/2010/to-trend-or-not-to-trend.html
- Twitter. (2014). *The Twitter government and elections handbook: 2014 U.S. edition*. Abgerufen von <https://archive.org/details/TwitterGovElectionsHandbook>
- Twitter. (2016). *How #Election2016 was tweeted so far*. Abgerufen von https://blog.twitter.com/en_us/a/2016/how-election2016-was-tweeted-so-far.html
- Twitter. (2018). *Update on Twitter's review of the 2016 US election*. Twitter Blog. Abgerufen von https://blog.twitter.com/official/en_us/topics/company/2018/2016-election-update.html
- Twitter. (2020). *Glossar*. Abgerufen von <https://help.twitter.com/de/glossary>
- Twitter. (2021). *Trend takeover and trend takeover+*. Abgerufen von <https://business.twitter.com/en/advertising/takeover/trend-takeovers.html>
- Van Aelst, P., van Erkel, P., D'heer, E., & Harder, R. A. (2017). Who is leading the campaign charts? Comparing individual popularity on old and new media. *Information, Communication & Society*, 20(5), 715–732. <https://doi.org/10.1080/1369118X.2016.1203973>
- van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199970773.001.0001>
- van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197–208. <https://doi.org/10.24908/ss.v12i2.4776>
- van Dijck, J., & Poell, T. (2013). Understanding social media logic. *Media and Communication*, 1(1), 2–14. <http://doi.org/10.17645/mac.v1i1.70>
- van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press. <https://doi.org/10.1093/oso/9780190889760.001.0001>
- Van Witsen, A. (2020). How daily journalists use numbers and statistics: The case of global average temperature. *Journalism Practice*, 14(9), 1047–1065. <https://doi.org/10.1080/17512786.2019.1682944>
- Vancil, D. L., & Pendell, S. D. (1987). The myth of viewer-listener disagreement in the first Kennedy-Nixon debate. *Central States Speech Journal*, 38(1), 16–27. <https://doi.org/10.1080/10510978709368226>
- von Krogh, T., & Andersson, U. (2016). Reading patterns in print and online newspapers: The case of the Swedish local morning paper VLT and online news site vlt.se. *Digital Journalism*, 4(8), 1058–1072. <https://doi.org/10.1080/21670811.2015.1111158>
- Vu, H. T. (2014). The online audience as gatekeeper: The influence of reader metrics on news editorial selection. *Journalism*, 15(8), 1094–1110. <https://doi.org/10.1177/1464884913504259>
- Waldherr, A., Geise, S., & Katzenbach, C. (2019). Because technology matters: Theorizing interdependencies in computational communication science with actor-network theory. *International Journal of Communication*, 13, 3955–3975.

- Wasike, B. (2019). Gender, nonverbal communication, and televised debates: A case study analysis of Clinton and Trump's nonverbal language during the 2016 town hall debate. *International Journal of Communication*, 13, 251–276.
- Webster, J. G. (2014). *The marketplace of attention: How audiences take shape in a digital age*. MIT Press. <https://doi.org/10.7551/mitpress/9892.001.0001>
- Wehner, J. (2020). „Numerical Coupling“: Zum Verhältnis von Massenmedien und Internet. In S. Maasen & J.-H. Passoth (Hrsg.), *Soziologie des Digitalen – Digitale Soziologie? Soziale Welt, Sonderband 23* (S. 176–197). Nomos. <https://doi.org/10.5771/9783845295008-176>
- Wendelin, M., Engelmann, I., & Neubarth, J. (2014). Nachrichtenfaktoren und Themen in Nutzerrankings. Ein Vergleich der journalistischen Nachrichtenauswahl und der Selektionsentscheidungen des Publikums im Internet. *Medien & Kommunikationswissenschaft (M&K)*, 62(3), 439–458. <https://doi.org/10.5771/1615-634x-2014-3-439>
- Winneg, K., & Jamieson, K. H. (2017). Learning from the 2016 U.S. general election presidential debates. *American Behavioral Scientist*, 61(4), 362–378. <https://doi.org/10.1177/0002764217702770>
- Winter, S., Metzger, M. J., & Flanagin, A. J. (2016). Selective use of news cues: A multiple-motive perspective on information selection in social media environments. *Journal of Communication*, 66(4), 669–693. <https://doi.org/10.1111/jcom.12241>
- Woolley, M., & Howard, P. N. (Hrsg.) (2019). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press. <https://doi.org/10.1093/oso/9780190931407.001.0001>
- Wu, A. X., & Harsh, T. (2020). Platform enclosure of human behavior and its measurement: Using behavioral trace data against platform episteme. *New Media & Society*. <https://doi.org/10.1177/1461444820933547>
- Zamith, R. (2018). A computational approach for examining the comparability of “most-viewed lists” on online news sites. *Journalism & Mass Communication Quarterly*, 95(1), 122–141. <https://doi.org/10.1177/1077699017714223>
- Zheng, P., & Shahin, S. (2020). Live tweeting live debates: How Twitter reflects and refracts the US political climate in a campaign season. *Information, Communication & Society*, 23(3), 337–357. <https://doi.org/10.1080/1369118X.2018.1503697>

EXTENDED ABSTRACT

Caught up in the rhetoric of technology, data, and media companies?
Problematic practices in dealing with nonscientific usage data

Silke Füst

Silke Fürst (M.A.), IKMZ – Department of Communication and Media Research, University of Zurich, Andreasstrasse 15, 8050 Zurich, Switzerland; Contact: s.fuerst(at)ikmz.uzh.ch.
ORCID: <https://orcid.org/0000-0002-1978-4277>

EXTENDED ABSTRACT

Caught up in the rhetoric of technology, data, and media companies?

Problematic practices in dealing with nonscientific usage data

Silke Füst

1. Introduction

Communication and media researchers, as well as scholars from related disciplines, are increasingly asking how social practices are measured, analyzed, and represented through data or digital traces (Bossetta, 2018; Dencik, 2020; Heintz, 2010; Mau, 2019; van Dijck & Poell, 2013; Wu & Harsh, 2020). Data increasingly permeate everyday life “and [have] become part of the social world” (Beer, 2018, p. 467). A growing body of literature analyzes how usage data influence the production, presentation, and use of media content. However, there is little research on the various ways in which media and communication research itself uses such nonscientific usage data. This paper shows that usage data from technology, data, market research, and media companies are in many cases opaque and often insufficiently scrutinized and contextualized by scholars. This is not only important in the context of digitization and big data but also concerns usage figures regarding traditional media. There is too little critical discussion and analysis as to what different metrics and rankings reveal and what commercial interests and strategies are driving them.

The analysis at hand identifies and discusses five problematic ways of dealing with nonscientific usage data: A) lacking contextualization of what usage metrics and rankings indicate, B) lacking contextualization of rankings as strategic business tools, C) unsubstantiated inference from usage data to macro phenomena, D) uncritical adoption of superlatives and generalizations created by vested interests, and E) use of estimated numbers as measured data.

2. The meaning of numbers in the context of change of science and media

Sociological works have highlighted that the reliance on numbers “as transparent and objective measurements of the ‘real’ world” (Christin, 2020, p. 117) has a long history. Since official statistics were initiated in the 18th century, data have become increasingly established as a form of societal self-description. Although data must be considered and interpreted in the context of their creation (boyd &

Crawford, 2012), they are attributed with a specific cultural legitimization: Numbers are commonly regarded as neutral and precise and are often associated with objectivity and scientificity (Beer, 2016; Heintz, 2010; Mau, 2019). Similarly, big data has been discursively linked to the myth that it would provide an objective and neutral access to complex realities (boyd & Crawford, 2012, p. 663). Companies offering digital data analytics emphasize that they can reflect reality in a comprehensive and precise manner (Beer, 2018). With digitization and the market power of big technology companies (van Dijck, Poell, & de Waal, 2018), usage data and rankings are now omnipresent in public communication and influence what becomes particularly visible and appears significant. Numerical comparisons (Heintz, 2010) are used in the competition for attention (Webster, 2014), imply winners and losers (Beer, 2016, pp. 17–26; Fürst, 2021) and the degree of publicity and social significance of content (Fürst & Oehmer, 2021; Gillespie, 2016, p. 67; Mau, 2019).

Moreover, usage figures communicated by commercial audience research and media, technology, and data companies have been such a common part of public discourse for many years that even scholars generally assume that these figures are based on measurements and indicate what they purport to. However, this paper shows that many usage numbers are used whose origin is nontransparent and whose meaning is unclear, or which, contrary to first appearances, are not based on a measurement process.

A critical examination of how science deals with commercial usage data is therefore important. The increase of publication pressure in science (Glerup & Horst, 2014) and data-driven “quick hit studies” (Lewis & Molyneux, 2018, p. 18) can undermine established scientific quality standards in the long term. Citing and incorporating numbers that cannot be verified and/or whose creation is opaque violates basic methodological rules and procedures in the social sciences as well as standards of research ethics (Brosius, Haas, & Koschel, 2016; Filipović, Klaus, & Strippel, 2016; McNabb, 2021). Therefore, it is important to identify and critique problematic practices in the handling of nonscientific usage data.

3. Problematic practices in dealing with nonscientific usage data

This paper analyzes five basic problematic practices (A–E) in the handling of nonscientific usage data and discusses them by giving examples. Problematic practices mean that the requirements of reliability and validity (Brosius et al., 2016; McNabb, 2021) are not sufficiently met or cannot be verified.

A. Lacking contextualization of what usage metrics and rankings indicate

With the emergence of social media platforms and online news sites, new usage metrics (e.g., followers, likes, shares) and rankings (most-read articles) became common. Given their strategic origin and economic importance (Bossetta, 2018; van Dijck, 2013, 2014), it is problematic that previous research has done comparatively little to examine the platforms’ infrastructures and the narratives used by technology, media, and data companies to promote their usage data (Beer,

2018, p. 466; Bossetta, 2018, p. 472; Lewis & Molyneux, 2018, p. 18; Turow & Couldry, 2018, p. 419). Social media analyses often hardly clarified or questioned what those metrics labeled as followers, likes, shares, etc. exactly mean and could indicate (Freelon, 2014). Moreover, it is problematic to interpret usage data and rankings generally as digital traces of users and use them to determine the popularity of actors (see, e.g., Francia, 2018; Van Aelst, van Erkel, D’heer, & Harder, 2017). The more that certain metrics are perceived as important by various publics, the greater the increase in attempts to influence and manipulate these metrics via payments, synchronization of numerous accounts, and automated processes (Fürst, 2021; Gorwa & Guilbeault, 2020; Turow & Couldry, 2018, p. 418).

B. Lacking contextualization of rankings as strategic business tools

Social media companies such as Twitter have developed various usage rankings to stimulate activity on their platform and gain media coverage (Fürst, 2021; Kreiss & McGregor, 2018, p. 165). Although these data obviously violate scientific standards due to their nontransparency and lack of validity, they are used as usage data in scientific papers. Twitter data on the “share of conversation” (Chadwick & Stromer-Galley, 2016, p. 284), “most tweeted topics” (Schill & Kirk, 2015, p. 215), and “top tweeted moments” (Schroeder, 2016, pp. 280–281) are cited, extensively reported, or even graphically depicted to provide empirical evidence for a particular argument or to illuminate an event from the users’ perspective. In doing so, studies adopted the narratives that Twitter has developed to be perceived as a barometer of public opinion, for example, by stating: “Twitter has allowed a much broader viewertariat to judge the debate’s defining moments” (Freelon & Karpf, 2015, p. 394). This problem also applies to many studies that analyze the Twitter trending topics. On the one hand, Twitter (2020) stated that these trends represent the “most popular topics on Twitter at any given time.” On the other hand, the company admitted that trends are not the topics that are most tweeted about (Twitter, 2010). The strong, but also constant and longer-term use of a hashtag promises less chances of placement in the trending topics than the strong and short-term spread of a new hashtag (Gillespie, 2011; Proferes & Summers, 2019). Twitter does not disclose more detailed information about the algorithm (Gillespie, 2011; Wu & Harsh, 2020). It is also not transparent on how many tweets the respective trending topics are based, how this compares to the total activity on Twitter (Bruns, 2020, pp. 75–76), and how many users are active on a daily basis (Bossetta, 2018, p. 481). However, many studies adopt the narrative set by Twitter and describe trending topics as the “most popular issues in the community” or “list of the most tweeted topics” (Marchetti & Ceccobelli, 2016) and thereby justify their research design.

C. Unsubstantiated inference from usage data to macro phenomena

Numerous studies have used the metrics of big technology companies to conduct big data analyses and draw conclusions about macro phenomena. For example, data on likes, retweets, and followers have been analyzed to infer the dominant

public opinion, which actors enjoy the greatest political support, or which candidates have the best electoral chances (for a critical overview Jungherr, 2016; Jungherr, Schoen, Posegga, & Jürgens, 2017; Lewis & Molyneux, 2018, pp. 16–17). Most often, such studies do not conceptualize this inference from usage data to the macro phenomenon of interest (Jungherr et al., 2017, p. 339). These largely unsubstantiated inferences also neglect the well-known fact that Twitter is not representative of the population because only a small proportion of citizens use Twitter, of whom, in turn, very few are highly active (boyd & Crawford, 2012, p. 669; Jungherr, 2016). Similar patterns exist in research on the role and influence of online usage data in journalism. Many studies that use rankings of “most read” articles for data analyses draw conclusions to the “content preferences among the public” (Boczkowski & Mitchelstein, 2013, p. 26) or the “audience desires” and “audience interests” (Lee, Lewis, & Powers, 2014, pp. 519–520), without substantiating this inference and in contradiction to findings of reception studies (Groot Kormelink & Costera Meijer, 2018; Webster, 2014, pp. 87–88).

D. Uncritical adoption of superlatives and generalizations created by vested interests

Technology, data, market research, and media companies often communicate usage data in terms of superlatives and generalizations. Some studies adopt them to emphasize the relevance of the event under study. For example, it has been postulated that televised debates in the context of the 2016 U.S. presidential election led to the “most tweeted day” (Camaj & Northup, 2019, p. 34) or a Twitter “record” (Jennings, Bramlett, McKinney, & Hardy, 2020, p. 1). Such superlatives are also found in studies examining the role of Twitter in the previous election cycle (Houston, McKinney, Hawthorne, & Spialek, 2013, p. 549). What is striking about this is that these statements and terms are based on media reports or PR statements from Twitter. Little or no data are provided, with nontransparent collection methods and dubious validity (e.g.: How were the tweets identified as referring to a specific event?). In addition, the connotation of the claimed superlatives varies, so that it remains unclear what the respective benchmarks and considered events are (see, for example, Fig. 1, “political-event record”). Thus, the claimed exceptionality signaled by such superlatives is less a data-based fact than a rhetorical strategy. By means of superlatives, the Twitter company postulates a continuous increase in the significance of its platform and is the only instance that has access to the data used for these statements. Therefore, it is important that researchers question such statements and disclose them as a PR strategy. This is similar to generalizations such as “Facebook election” (Johnson & Perlmutter, 2011) or “Twitter election” (Houston et al., 2013) and recurring assertions of record viewership of televised debates (Winneg & Jamieson, 2017, pp. 362, 374, 376–377). The latter originates from the media releases of market research and media companies and are already proclaimed and widely disseminated during the run-up to events (Fürst, 2021). Critical analyses of such superlatives and generalizations are important and can reveal that television viewership—rather than

breaking records—tends to be in long-term decline (Holz, Akin, & Jamieson, 2016; Prior & Bougher, 2018, pp. 830–831).

E. Use of estimated numbers as measured data

Finally, there are also numbers that are based only on estimations but are adopted by research as if they were empirical statements or measured data. Taking the example of studies on the diffusion of the Internet: There are no data on the worldwide number of Internet users in the 1990s (Hargittai, 1999; Hoffman & Novak, 1994), and yet such numbers have been presented in scientific studies. These are mostly based on the statements of market research and consulting companies. Such companies repeatedly published high and rapidly rising figures for worldwide Internet use, feeding euphoric expectations of the technology and stimulating investments in the Internet sector (Jordan, 2001; Latzer, 1997, p. 41). These figures often differed greatly and can be described as “dubious” (Latzer, 1997, p. 41) because of their lacking transparency regarding their origin. In some cases, they were derived from the number of hosts multiplied by the guessed number of users per host (Fürst, 2017, p. 49): “What happens when you multiply the number of hosts by the number of users per host? Depending on which numbers you select, you get something between 2.5 million and 32 million users” (Hoffman & Novak, 1994). Nevertheless, figures on the worldwide use of the Internet in the 1990s were cited in communication and media studies. Some works give such dubious figures to illustrate the relevance of their object of study (Bobbitt, 1995, p. 27), to trace a historical development (Faulstich, 2012, p. 399), or to contribute evidence for a particular perspective: “With 25 million people estimated to be communicating on the Internet, should communication researchers now consider this network of networks a mass medium?” (Morris & Ogan, 1996, p. 39). Other studies give numbers regarding the Internet use in the early and mid-1990s to show diffusion patterns and relied on the information given by companies and organizations such as NUA, NVA Ltd., and the International Telecommunication Union, which did not provide information on how they arrived at the numbers. Scientists used these figures for data analysis or created graphs without reflecting on the source and validity of the numbers (e.g., Rogers, 2003, p. 347).

A similar problem exists regarding widely disseminated television viewership numbers for global media events. Although there are no measured audience ratings on a global scale (Bourdon & Méadel, 2014), communication and media studies rely on global viewer numbers for conceptual decisions and for demonstrating the relevance of the event under study (e.g., Dayan & Katz, 1992; Solves, Sánchez, & Rius, 2018). These numbers stem from news media reports, PR messages of involved actors, and estimated numbers of commercial audience research and are also published before the concerned media events have taken place (Fürst, 2020).

4. Conclusion

Studies in communication and media research increasingly rely on nonscientific usage data from technology, data, market research, and media companies. Critical scholars have raised concerns on specific points in certain areas of research. However, so far, there has been no broad and comprehensive discussion of how to deal with nonscientific usage data and how to avoid problematic practices. Based on numerous literature references and with the inclusion of various areas of communication and media research, this analysis contributes to gaining an overview and pinpoints several significant problematic practices in dealing with nonscientific usage data. These were often located in quantitative research on social media, political communication, media use, and journalism. However, such practices have also been noted in studies in media science, media history, and media sociology (see especially Section E), and thus affect broad sections of the field.

The usage figures discussed here (A–E) are coupled with nontransparency, strategic communication, and / or problems of validity. However, studies often—and mostly without critical reflection—resort to such figures to justify the relevance of the research object or research question, to lend empirical evidence to certain arguments, conceptual decisions, or derived developments, to emphasize the relevance of their research results, and to draw conclusions on macro phenomena. In some cases, this involves only a brief presentation of a few numbers; in others, many figures are included or displayed graphically. Moreover, nonscientific usage data influence data selection and collection as well as the research design or are even used for data analysis. Thus, the identified problematic practices encompass practically all steps of scientific knowledge production. However, this study cannot show how much of the total research is affected by such practices. It also does not claim to have covered all conceivable problematic practices regarding nonscientific usage data. Many scientists have probably not used or cited such data. However, this is not a sufficient mechanism for ensuring the quality of scientific knowledge production. Studies with problematic practices are cited by other works and have become part of the scientific knowledge. Concealed by citation chains, the numbers given by them are eventually passed on without awareness of the problematic origin (see, e.g., Fürst, 2020, p. 1526).

This paper not only sheds a critical light on communication and media research, the partial lack of ensuring scientific quality, and recent developments in science but also opens up new questions and perspectives with which communication and media research can make an important contribution. A critical look at usage data is essential, especially in societies in which usage figures and rankings are omnipresent and powerful corporations largely control access to data. Communication and media researchers could critically examine widely disseminated usage numbers, discuss them in light of data gathered by various methodological approaches, and reveal the narratives, interests, and strategies of the companies and organizations involved. Nonscientific usage numbers are evidently often disseminated to influence media usage, evoke certain imaginations of the co-audience, and trigger self-fulfilling prophecies (Fürst, 2017, 2020, 2021; Webster, 2014). Communication and media research can stimulate more reflection in soci-

ety and public discourse on how people are represented by figures and how attention dynamics are stimulated by means of numbers. This would be a contribution to strengthening the critical potential of communication and media research, which has been increasingly called for in recent years (Dencik, 2020; Fuchs, 2017; Lewis & Molyneux, 2018).

References

- Beer, D. (2016). *Metric power*. Palgrave Macmillan. <https://doi.org/10.1057/978-1-137-55649-3>
- Beer, D. (2018). Envisioning the power of data analytics. *Information, Communication & Society*, 21(3), 465–479. <https://doi.org/10.1080/1369118X.2017.1289232>
- Bobbitt, R. (1995). An Internet primer for public relations. *Public Relations Quarterly*, 40(3), 27–32.
- Boczkowski, P. J., & Mitchelstein, E. (2013). *The news gap: When the information preferences of the media and the public diverge*. MIT Press.
- Bossetta, M. (2018). The digital architectures of social media: Comparing political campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 U.S. Election. *Journalism & Mass Communication Quarterly*, 95(2), 471–496. <https://doi.org/10.1177/1077699018763307>
- Bourdon, J., & Méadel, C. (Eds.). (2014). *Television audiences across the world: Deconstructing the ratings machine*. Palgrave Macmillan. <https://doi.org/10.1057/9781137345103>
- boyd, d., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Brosius, H.-B., Haas, A., & Koschel, F. (2016). *Methoden der empirischen Kommunikationsforschung. Eine Einführung* [Methods of empirical communication research: An introduction] (7th ed.). Springer VS. <https://doi.org/10.1007/978-3-531-19996-2>
- Bruns, A. (2020). Big social data approaches in Internet studies: The case of Twitter. In J. Hunsinger, M. M. Allen, & L. Klastrup (Eds.), *Second international handbook of Internet research* (pp. 65–81). Dordrecht: Springer. https://doi.org/10.1007/978-94-024-1555-1_3
- Camaj, L., & Northup, T. (2019). Dual-screening the candidate image during presidential debates: The moderating role of Twitter and need to evaluate for the effects on candidate perceptions. *Journal of Broadcasting & Electronic Media*, 63(1), 20–38. <https://doi.org/10.1080/08838151.2019.1574117>
- Chadwick, A., & Stromer-Galley, J. (2016). Digital media, power, and democracy in parties and election campaigns: Party decline or party renewal? *The International Journal of Press/Politics*, 21(3), 283–293. <https://doi.org/10.1177/1940161216646731>
- Christin, A. (2020). What data can do: A typology of mechanisms. *International Journal of Communication*, 14, 1115–1134.
- Dayan, D., & Katz, E. (1992). *Media events: The live broadcasting of history*. Cambridge: Harvard University Press.
- Dencik, L. (2020). Mobilizing media studies in an age of datafication. *Television & New Media*, 21(6), 568–573. <https://doi.org/10.1177/1527476420918848>

- Faulstich, W. (2012). *Die Mediengeschichte des 20. Jahrhunderts* [Media history of the 20th century]. Wilhelm Fink.
- Filipović, A., Klaus, E., & Strippel, C. (2016). Der Ethik-Kodex der DGPK [The code of ethics of the German association DGPK]. *Medien & Kommunikationswissenschaft (M&K)*, 64(1), 93–108. <https://doi.org/10.5771/1615-634X-2016-1-93>
- Francia, P. L. (2018). Free media and Twitter in the 2016 presidential election: The unconventional campaign of Donald Trump. *Social Science Computer Review*, 36(4), 440–455. <https://doi.org/10.1177/0894439317730302>
- Freelon, D. (2014). On the interpretation of digital trace data in communication and social computing research. *Journal of Broadcasting & Electronic Media*, 58(1), 59–75. <https://doi.org/10.1080/08838151.2013.875018>
- Freelon, D., & Karpf, D. (2015). Of big birds and bayonets: Hybrid Twitter interactivity in the 2012 Presidential debates. *Information, Communication & Society*, 18(4), 390–406. <https://doi.org/10.1080/1369118X.2014.952659>
- Fuchs, C. (2017). *Social media: A critical introduction* (2nd ed.). Sage.
- Fürst, S. (2017). Die Etablierung des Internets als Self-Fulfilling Prophecy? Zur Rolle der öffentlichen Kommunikation bei der Diffusion neuer Medien [Popularization of the Internet as a self-fulfilling prophecy? The role of public communication in the diffusion of new media]. *Medien & Zeit*, 32(2), 43–55.
- Fürst, S. (2020). “The whole world watching”? How news media create the myth of an audience of billions and foster imagined communities. *International Journal of Communication*, 14, 1524–1541.
- Fürst, S. (2021). Neue Öffentlichkeitsdynamiken: Zu selbstverstärkenden, plattformübergreifenden Effekten von ‚Popularität‘ [New dynamics of the public sphere: Self-reinforcing effects of ‘popularity’ across platforms]. In M. Eisenegger, M. Prinzing, P. Ettinger, & R. Blum (Eds.), *Digitaler Strukturwandel der Öffentlichkeit. Historische Verortung, Modelle und Konsequenzen* [Digital structural change of the public sphere: Historical perspectives, models, and consequences] (pp. 339–359). Springer VS. https://doi.org/10.1007/978-3-658-32133-8_19
- Fürst, S., & Oehmer, F. (2021). Attention for attention hotspots: Exploring the newsworthiness of public response in the metric society. *Journalism Studies*, 22(6), 799–819. <https://doi.org/10.1080/1461670X.2021.1889396>
- Gillespie, T. (2011). Can an algorithm be wrong? Twitter Trends, the specter of censorship, and our faith in the algorithms around us. Retrieved from <http://culturedigitally.org/2011/10/can-an-algorithm-be-wrong/>
- Gillespie, T. (2016). #Trendingistrending: When algorithms become culture. In R. Seyfert & J. Roberge (Eds.), *Algorithmic cultures: Essays on meaning, performance and new technologies* (pp. 52–75). Routledge.
- Glerup, C., & Horst, M. (2014). Mapping ‘social responsibility’ in science. *Journal of Responsible Innovation*, 1(1), 31–50. <https://doi.org/10.1080/23299460.2014.882077>
- Gorwa, R., & Guilbeault, D. (2020). Unpacking the social media bot: A typology to guide research and policy. *Policy & Internet*, 12(2), 225–248. <https://doi.org/10.1002/poi3.184>
- Groot Kormelink, T., & Costera Meijer, I. (2018). What clicks actually mean: Exploring digital news user practices. *Journalism*, 19(5), 668–683. <https://doi.org/10.1177/1464884916688290>

- Hargittai, E. (1999). Weaving the western web: Explaining differences in Internet connectivity among OECD countries. *Telecommunications Policy*, 23, 701–718. [https://doi.org/10.1016/S0308-5961\(99\)00050-6](https://doi.org/10.1016/S0308-5961(99)00050-6)
- Heintz, B. (2010). Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs [Numerical difference: Toward a sociology of (quantitative) comparisons]. *Zeitschrift für Soziologie*, 39(3), 162–181. <https://doi.org/10.1515/zfsoz-2010-0301>
- Hoffman, D. L., & Novak, T. P. (1994). *Wanted: Net.Census*. Retrieved from <https://www.wired.com/1994/11/hoffman-if/>
- Holz, J., Akin, H., & Jamieson, K. H. (2016). *Presidential debates: What's behind the numbers?* Retrieved from <https://www.annenbergpublicpolicycenter.org/publication/presidential-debates-whats-behind-the-numbers/>
- Houston, J. B., McKinney, M. S., Hawthorne, J., & Spialek, M. L. (2013). Frequency of tweeting during presidential debates: Effect on debate attitudes and knowledge. *Communication Studies*, 64(5), 548–560. <https://doi.org/10.1080/10510974.2013.832693>
- Jennings, F. J., Bramlett, J. C., McKinney, M. S., & Hardy, M. M. (2020). Tweeting along partisan lines: Identity-motivated elaboration and presidential debates. *Social Media + Society*, 6(4), 1–12. <https://doi.org/10.1177/2056305120965518>
- Johnson, T. J., & Perlmutter, D. D. (Eds.). (2011). *New media, campaigning and the 2008 Facebook election*. Routledge. <https://doi.org/10.4324/9781315872650>
- Jordan, T. (2001). Measuring the Internet: Host counts versus business plans. *Information, Communication & Society*, 4(1), 34–53. <https://doi.org/10.1080/13691180122452>
- Jungherr, A. (2016). Twitter use in election campaigns: A systematic literature review. *Journal of Information Technology & Politics*, 13(1), 72–91. <https://doi.org/10.1080/19331681.2015.1132401>
- Jungherr, A., Schoen, H., Posegga, O., & Jürgens, P. (2017). Digital trace data in the study of public opinion: An indicator of attention toward politics rather than political support. *Social Science Computer Review*, 35(3), 336–356. <https://doi.org/10.1177/0894439316631043>
- Kreiss, D., & McGregor, S. C. (2018). Technology firms shape political communication: The work of Microsoft, Facebook, Twitter, and Google with campaigns during the 2016 U.S. presidential cycle. *Political Communication*, 35(2), 155–177. <https://doi.org/10.1080/10584609.2017.1364814>
- Latzer, M. (1997). *Mediamatik – Die Konvergenz von Telekommunikation, Computer und Rundfunk* [Mediamatics – The convergence of telecommunications, computers, and broadcasting]. Westdeutscher Verlag. <https://doi.org/10.1007/978-3-322-86878-7>
- Lee, A. M., Lewis, S. C., & Powers, M. (2014). Audience clicks and news placement: A study of time-lagged influence in online journalism. *Communication Research*, 41(4), 505–530. <https://doi.org/10.1177/0093650212467031>
- Lewis, S. C., & Molyneux, L. (2018). A decade of research on social media and journalism: Assumptions, blind spots, and a way forward. *Media and Communication*, 6(4), 11–23. <http://doi.org/10.17645/mac.v6i4.1562>
- Marchetti, R., & Ceccobelli, D. (2016). Twitter and television in a hybrid media system: The 2013 Italian election campaign. *Journalism Practice*, 10(5), 626–644. <https://doi.org/10.1080/17512786.2015.1040051>
- Mau, S. (2019). *The metric society: On the quantification of the social*. Polity.

- McNabb, D. E. (2021). *Research methods for political science: Quantitative, qualitative and mixed method approaches* (3rd ed.). Routledge. <https://doi.org/10.4324/9781003103141>
- Morris, M., & Ogan, C. (1996). The Internet as mass medium. *Journal of Communication*, 46(1), 39–50. <https://doi.org/10.1111/j.1460-2466.1996.tb01460.x>
- Prior, M., & Bougher, L. D. (2018). “Like they’ve never, ever seen in this country”? Political interest and voter engagement in 2016. *Public Opinion Quarterly*, 82(S1), 822–842. <https://doi.org/10.1093/poq/nfy002>
- Proferes, N., & Summers, E. (2019). Algorithms and agenda-setting in Wikileaks’ #Podestae-mails release. *Information, Communication & Society*, 22(11), 1630–1645. <https://doi.org/10.1177/2056305117698493>
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
- Schill, D., & Kirk, R. (2015). Issue debates in 140 characters: Online talk surrounding the 2012 debates. In J. A. Hendricks (Ed.), *Presidential campaigning and social media: An analysis of the 2012 campaign* (pp. 198–217). Oxford University Press.
- Schroeder, A. (2016). *Presidential debates: Risky business on the campaign trail* (3rd ed.). Columbia University Press. <http://doi.org/10.7312/columbia/9780231170574.001.0001>
- Solves, J., Sánchez, S., & Rius, I. (2018). The prince and the pauper: Journalistic culture and Paralympic Games in the Spanish print press. *Journalism*, 19(12), 1713–1728. <https://doi.org/10.1177/1464884916671894>
- Turow, J., & Couldry, N. (2018). Media as data extraction: Towards a new map of a transformed communications field. *Journal of Communication*, 68(2), 415–423. <https://doi.org/10.1093/joc/jqx011>
- Twitter. (2010). *To trend or not to trend*. Retrieved from https://blog.twitter.com/en_us/a/2010/to-trend-or-not-to-trend.html
- Twitter. (2020). *Glossar*. Retrieved from <https://help.twitter.com/de/glossary>
- Van Aelst, P., van Erkel, P., D’heer, E., & Harder, R. A. (2017). Who is leading the campaign charts? Comparing individual popularity on old and new media. *Information, Communication & Society*, 20(5), 715–732. <https://doi.org/10.1080/1369118X.2016.1203973>
- van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199970773.001.0001>
- van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197–208. <https://doi.org/10.24908/ss.v12i2.4776>
- van Dijck, J., & Poell, T. (2013). Understanding social media logic. *Media and Communication*, 1(1), 2–14. <http://doi.org/10.17645/mac.v1i1.70>
- van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press. <https://doi.org/10.1093/oso/9780190889760.001.0001>
- Webster, J. G. (2014). *The marketplace of attention: How audiences take shape in a digital age*. MIT Press. <https://doi.org/10.7551/mitpress/9892.001.0001>
- Winneg, K., & Jamieson, K. H. (2017). Learning from the 2016 U.S. general election presidential debates. *American Behavioral Scientist*, 61(4), 362–378. <https://doi.org/10.1177/0002764217702770>
- Wu, A. X., & Harsh, T. (2020). Platform enclosure of human behavior and its measurement: Using behavioral trace data against platform episteme. *New Media & Society*. <https://www.doi.org/10.1177/1461444820933547>